

TI 2011-148/1

Tinbergen Institute Discussion Paper



An Optimal Signaling Equilibrium

Vladimir Karamychev

Bauke Visser

Erasmus School of Economics, Erasmus University Rotterdam, Tinbergen Institute.

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at <http://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Duisenberg school of finance is a collaboration of the Dutch financial sector and universities, with the ambition to support innovative research and offer top quality academic education in core areas of finance.

DSF research papers can be downloaded at: <http://www.dsf.nl/>

Duisenberg school of finance
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 8579

An Optimal Signaling Equilibrium

Vladimir Karamychev and Bauke Visser

Erasmus University Rotterdam and Tinbergen Institute

This version: October 2011.

Abstract. This paper analyses the optimal combination of costly and costless messages that a Sender uses in a signaling game if he is able to choose among all equilibrium communication strategies. We provide a complete characterization of the equilibrium that maximizes the Sender's *ex ante* expected utility in case of uniformly distributed types and quadratic loss functions. First, the Sender often wants to avoid money burning by using the most informative cheap talk communication strategy. Second, if he does burn money, he avoids separation and only re-arranges the existing intervals of the most informative cheap talk equilibrium, possibly adding one extra interval. Money burning takes place in the second interval only.

Key Words: Cheap talk, money burning, optimal equilibrium.

JEL Classification: C 7, D 8.

1. Introduction

Communication in signaling games can take place through costly and costless signals. Crawford and Sobel (1982) present what has become the canonical model of communication through costless signals - cheap talk - between a sender who is privately informed about the state and an uninformed receiver who has to take an action that affects the utility of both players.¹ They show that cheap talk can be informative as long as the degree to which the preferences of the two players concerning the best action in a given state do not differ too much. Although informative, communication will necessarily be vague in the sense that the Sender adds noise to his messages: equilibrium communication strategies partition the (single-dimensional) type space in intervals, and the Sender only reports in which element of the partition his type lies.

Costly signals were first studied in their extreme form of completely dissipative signals - money burning - by Spence (1973). In the context of labor market signaling, he shows how high ability job applicants could separate themselves from low ability ones by showing credentials that are costly to obtain but otherwise useless. This precision of communication through money burning contrasts with the vagueness of communication based on cheap talk.

In practice, communication often relies on a combination of cheap talk and costly signals (dissipative or not). This situation is studied by Austen-Smith and Banks (1998, 2000). They show that the option to burn money expands the set of equilibria enormously, and can increase the precision of cheap talk communication.² In this paper, we follow the same approach by allowing the Sender to use both cheap talk and money burning, and characterize all possible equilibrium communication strategies. Our main contribution is the derivation of

¹ See also Green and Stokey (2007) for an earlier and an alternative formulation.

² Austen-Smith and Banks (1998) study some welfare results for the uniform-quadratic case. Kartik (2007) shows that cheap talk can be influential with money burning if and only if it can be influential without money burning.

the optimal Sender's equilibrium, *i.e.*, the communication strategy that maximizes the Sender's expected utility *ex ante*, before knowing his type.

We study this optimal Sender's equilibrium for the following reasons. First, from a theoretical point of view it is interesting to know the optimal mixture of costly and costless signals from the Sender's prospective. Since it is up to the Sender whether he uses costly signaling in addition to cheap talk, we want to know whether he wants to use money burning or not and, if yes, to which extent.

Second, this is a natural focal point for the Sender if he can choose an equilibrium. This focal point is plausible because the Sender is the first mover in the signaling game and, therefore, might have an advantage to announce which equilibrium he will be playing. Finally, the optimal Sender's equilibrium might have nice evolutionary stability properties, as the recent work of Demichelis and Weibull (2008) demonstrates.

In the main part of the paper, we augment the much-used uniform-quadratic variant of the Crawford and Sobel (1982) cheap talk model with costly signaling (money burning). In this model, the Sender's type is uniformly distributed, and the players' preferences are represented by quadratic loss functions.

We find that the optimal equilibrium is similar to the most informative equilibrium of Crawford and Sobel (1982). Depending on the preference misalignment parameter of the model, the Sender either burns money or not. In particular, if the degree of misalignment exceeds a threshold value, the latter case realizes and the optimal equilibrium is just the most informative pure cheap talk equilibrium. In this case, burning money does not allow the Sender to improve his pay-off because it is too costly.

Only when the preference misalignment parameter is smaller than the threshold value, and satisfies some restrictions, does money burning occur in equilibrium. Nevertheless, the Sender never uses it to separate the types perfectly. The optimal equilibrium is again a partition equilibrium very similar to the most informative pure cheap talk equilibrium.

The Sender uses money burning in two ways. First, he may increase the number of intervals in the equilibrium partition. In this case, the optimal equilibrium is obtained from the most informative cheap talk equilibrium by squeezing the existing partition and inserting an extra interval in between the first and the second ones. All types from the new (second) interval pool together, send identical message, and burn the same amount of money; no other types use costly signaling. Although he burns a positive amount of money, the Sender's expected utility increases because the resulting partition becomes finer.

Second, even when having more intervals is not optimal (too costly), the Sender burns money in order to adjust the sizes of the existing intervals to his benefit. In this case, the optimal equilibrium is obtained from the most informative cheap talk equilibrium by squeezing the second and stretching all the other intervals. All types from the second interval pool together, send identical message, and burn a positive amount of money. The expected utility of the sender increases because intervals in the resulting partition become more even.

As to the values of the misalignment parameter for which the Sender burns money in the optimal equilibrium, we show that money burning can occur only when the degree of misalignment between Sender's and Receiver's preferences is such that the corresponding cheap-talk equilibrium has at least three intervals. Moreover, money burning is only used when the value of the misalignment parameter is close to a value at which the maximum number of intervals changes in a pure cheap talk game. When the misalignment gradually and continuously vanishes, optimal equilibria with and without money burning form an infinitely alternating pattern and the Sender's *ex-ante* probability of sending a costly signal converges to zero.

We conclude our analysis by studying general specifications of the type distribution and objective function of the Sender. We derive a sufficient condition under which the separation of types is not optimal. We also show that for any distribution and any objective function,

there will be a type for which this condition holds. Consequently, perfectly separating equilibria are never optimal.

Recently, De Haan *et al* (2011) have studied how subjects in a laboratory experiment use cheap talk and money burning in the uniform-quadratic case for various degrees of preference misalignment. It is unlikely that any lab experiment will be able to mimic the precise alternating pattern of the presence and absence of money burning as a function of the misalignment parameter that we have shown to be the best from the Sender's perspective. They do find, however, that Senders have a strong preference for cheap talk, and that only when interests start to grow more misaligned do Senders turn to money burning. They also find, again in line with our theoretical observations, that Senders do not burn money to support intervals of full separation.

The rest of the paper is organized as follows. Section 2 sets up the original signaling model and derives all its WPBE. Section 3 derives the Sender's optimal equilibrium. Section 4 generalizes some of the results to general types' distributions and Senders' utility functions, Section 5 concludes, and Appendix contains all the proofs.

2. Signaling Game

In the main part of the analysis, we use a uniform-quadratic version of the signaling game from Austen-Smith and Banks (2000). The main reason to start the analysis in this way is the fact that this version of the pure cheap talk game has been used extensively in applied work on communication. It is therefore particularly important to understand how this version behaves when one adds money burning. Section 4 generalizes some of the results to general types' distributions and Senders' utility functions.

There are two players, a Sender S and a Receiver R . The type t of the Sender is his private information and is uniformly distributed over the type space $[0, 1]$. Having observed t the Sender sends the Receiver a costless message $m \in \mathbf{M}$ from any continuum space \mathbf{M} , and

publicly burns an amount of money $b \geq 0$. Let $\sigma : [0, 1] \rightarrow \mathbf{M} \times \mathfrak{R}_+$ be the pure signaling strategy of S, $\sigma = (m(t), b(t))$, and let set $\mathbf{T}(s, \sigma)$ be the pre-image of s under σ : it is a set of types that, by following the strategy $\sigma(t)$, send signal $s = (m, b) \in \mathbf{M} \times \mathfrak{R}_+$:

$$\mathbf{T}(s, \sigma) \equiv \sigma^{-1}(s) = \{t \in [0, 1] : \sigma(t) = s\}.$$

Having observed signal $s = (m, b)$, the Receiver forms his *posterior* beliefs about the type of the Sender, which we denote by the CDF $G(z|s) \equiv \Pr(t \leq z)$. Then, the Receiver takes an action $a \in \mathfrak{R}$. Let $\alpha : \mathbf{M} \times \mathfrak{R}_+ \rightarrow \mathfrak{R}$ be his pure strategy, $\alpha(s)$. By $a(t) \equiv \alpha(\sigma(s))$ we refer to the action induced by type t .

As is usual in cheap talk signaling games, the model is characterized by an enormous multiplicity of equilibria. That is why we consider classes of WPBE such that all equilibria which generate the same induced action function $a(t) \equiv \alpha(\sigma(s))$ belong to the same class. We refer to the whole class of these equilibria as to equilibrium.

Preferences of the Sender and the Receiver are given by the following utility functions U^S and U^R respectively:

$$U^S(a, t, s) \equiv U^S(a, t, (m, b)) = -(t + x - a)^2 - b,$$

and

$$U^R(a, t, s) \equiv U^R(a, t, (m, b)) = -(t - a)^2,$$

where $x > 0$ is the parameter indicating the extent to which the players' preferences over (a, t) -pairs differ. Players' *ex-ante* expected utilities from a strategy profile (σ, α) are given by:

$$V^S(\sigma, \alpha) \equiv E(U^S(\alpha(\sigma(t)), t, \sigma(t))) = -E((t + x - a(t))^2) - E(b(t)), \quad (1)$$

$$V^R(\sigma, \alpha) \equiv E(U^R(\alpha(\sigma(t)), t, \sigma(t))) = -E((t - a(t))^2). \quad (2)$$

The definition of a weak perfect Bayesian equilibrium (WPBE) $\Omega = \langle \sigma(t), \alpha(s), G(t|s) \rangle$ is standard and requires (i) $\sigma(t)$ to be optimal given $\alpha(s)$, $\alpha(s)$ to be optimal given $G(t|s)$, and for s where $\Pr(s) > 0$ (on equilibrium path) $G(t|s)$ to be consistent with $\sigma(t)$.

Austen-Smith and Banks (2000) show that all equilibria of the model are essentially partition equilibria where the type space is a disjoint union of the set of pooling types \mathbf{P} with the set of separating types \mathbf{S} , i.e., $[0, 1] = \mathbf{S} \cup \mathbf{P}$, $\mathbf{S} \cap \mathbf{P} = \emptyset$. The set \mathbf{S} consists of all types that send individual messages, which we assume to be

$$m^*(t) \equiv \text{“my type is } t\text{”},$$

burn distinct amounts of money $b^s(t)$, and elicit distinct actions. To the contrary, the set \mathbf{P} is the union of mutually disjoint pooling intervals. By neglecting marginal types, which measure is zero, we index the pooling intervals by a subscript $i \in \mathbf{I}$, where the size of the index set \mathbf{I} is at most countable infinity, and write them as (A_i, B_i) , $A_i < B_i$. We denote the set of such pooling intervals by $\mathbf{A} = \{(A_i, B_i) | i \in \mathbf{I}\}$. For any pooling interval $(A, B) \in \mathbf{A}$, all types $t \in (A, B)$ send identical messages, which we assume to be

$$m^*(A, B) \equiv \text{“my type is an element of } (A, B)\text{”},$$

burn identical amounts of money b_i^P , and elicit identical actions. Since \mathbf{A} uniquely determines the sets \mathbf{P} and \mathbf{S} by $\mathbf{P} = \bigcup_{(A_i, B_i) \in \mathbf{A}} (A_i, B_i)$ and $\mathbf{S} = [0, 1] \setminus \mathbf{P}$, we also refer to \mathbf{A} as to equilibrium partition.

In order to deal with strategies that have infinitely many pooling intervals we consider them as a limit of a strategy with finite number N of intervals when N unboundedly increases. For finite N , any partition \mathbf{A} can be written as follows:

$$\mathbf{A} = \{(A_i, B_i) | i = 1, \dots, N\}, \quad 0 \equiv B_0 \leq A_i < B_i \leq A_{i+1} \leq A_{N+1} \equiv 1, \quad i = 1, \dots, N, \quad (3)$$

where types $t \in (A_i, B_i)$ pool together and types $t \in (B_i, A_{i+1})$ separate. If $N = 0$ then $\mathbf{A} = \{\emptyset\}$, $A_1 = 1$, and $\mathbf{S} = [0, 1]$, *i.e.*, all types separate. We use $\mathbf{\Pi}(N)$ to denote the set of all partitions \mathbf{A} with up to N intervals. Finally, we denote the amount of money burned by pooling types $t \in (A_i, B_i)$ by $b(t) = b_i^P$, and by separating types $t \in (B_i, A_{i+1})$ by $b(t) = b^S(t)$.

According to Austen-Smith and Banks (2000) in any WPBE of the signaling game, the Sender partitions the type-space according to a finite partition $\mathbf{A} \in \mathbf{\Pi}(N)$ or to a limit thereof when N increases unboundedly. It turns out that the reverse is also true, *i.e.*, for any partition $\mathbf{A} \in \mathbf{\Pi}(N)$ there exists a corresponding WPBE. In order to characterize all such equilibria, we define a function c^P :

$$c^P(A, B, c) \equiv (B + A)x - (B - A)^2 / 4 + c. \quad (4)$$

The following proposition provides the equilibrium characterization result.

Proposition 1. For any partition $\mathbf{A} \in \mathbf{\Pi}(N)$, there exist a money burning parameter c and an off equilibrium path action $\alpha^{OFF} \in [0, 1]$ such that the tuple $\langle \mathbf{A}, c, \alpha^{OFF} \rangle$ defines a (class of) WPBE Ω of the game where:

- a) For any $i = 1, \dots, N$, a type $t \in (A_i, B_i)$ burns $b_i^P = c^P(A_i, B_i, c) \geq 0$ and sends message $m^*(A_i, B_i)$. Having observed the signal $s = (m^*(A_i, B_i), b_i^P)$, the Receiver sets $\alpha(s) = (A_i + B_i) / 2$. Type t Sender gets utility $u^\Omega(t) = -x^2 - 2xt - c + (B_i - t)(t - A_i)$.
- b) For any $i = 1, \dots, N$, a type $t \in (B_i, A_{i+1})$ burns $b^S(t) = 2xt + c \geq 0$ and sends messages $m^*(t)$. Having observed the signal $s = (m^*(t), b^S(t))$, the Receiver sets $\alpha(s) = t$. Type t Sender gets utility $u^\Omega(t) = -x^2 - 2xt - c$.
- c) Having observed any off equilibrium path signal s^{OFF} , the Receiver sets $\alpha(s^{OFF}) = \alpha^{OFF}$.
- d) Players get *ex-ante* expected utilities:

$$V^S(\Omega) = -c - x^2 - x + \frac{1}{6} \sum_{i=1}^N (B_i - A_i)^3 \text{ and } V^R(\Omega) = -\frac{1}{12} \sum_{i=1}^N (B_i - A_i)^3 \quad (5)$$

e) Every WPBE of the signaling game allows for such a representation.

Parameter c in a WPBE is an additive constant to the amounts of money that all types burn. For a given partition \mathbf{A} , its value is not necessarily unique. On the one hand, it cannot be too small as types cannot burn negative amounts, *i.e.*, $b_i^P = c^P(A_i, B_i, c) \geq 0$ and $b^S(t) = 2xt + c \geq 0$. This defines the lower bound $\underline{c}(\mathbf{A})$ as the lowest value of c for which all $b_i^P \geq 0$ and $b^S(t) \geq 0$, and at least one type burns no money. On the other hand, c cannot be too large either as otherwise no off equilibrium path beliefs support this WPBE. This defines the upper bound $\bar{c}(\mathbf{A})$ as the highest value of c for which this WPBE still exists. The following example demonstrates the general structure of an equilibrium and its multiplicity.

Example 1. Suppose $x = 0.15$ and consider partition $\mathbf{A} = \{(0, 0.2), (0.3, 1)\}$. One can check that for any $c \in [\underline{c}, \bar{c}] = [-0.02, 0]$ and off equilibrium path actions $\alpha^{OFF} \in [0, x - \sqrt{0.0225 + c}]$, $\langle \mathbf{A}, c, \alpha^{OFF} \rangle$ defines a WPBE such that types $t \in (0, 0.2)$ pool, types $t \in (0.2, 0.3)$ separate, and types $t \in (0.3, 1)$ pool.

In Figure 1, we have drawn the equilibrium money burning function $b^\Omega(t)$ (dotted bold line) and the Sender's utility (drawn bold line), in case of $c = \underline{c} = -0.02$. Types $t \in (0, 0.2)$ burn no money, types $t \in (0.2, 0.3)$ burn $b^S(t) = 0.3t - 0.02$, and types $t \in (0.3, 1)$ burn $b_3^P(t) = 0.0525$. This is the lowest value of the parameter c for which WPBE exists: for $c < \underline{c}$ types $t \in (0, 0.2)$ would need to burn a negative amount of money, which is not feasible. This equilibrium is unique up to off equilibrium path actions: any $\alpha^{OFF} \in [0, 0.1]$ supports this WPBE.

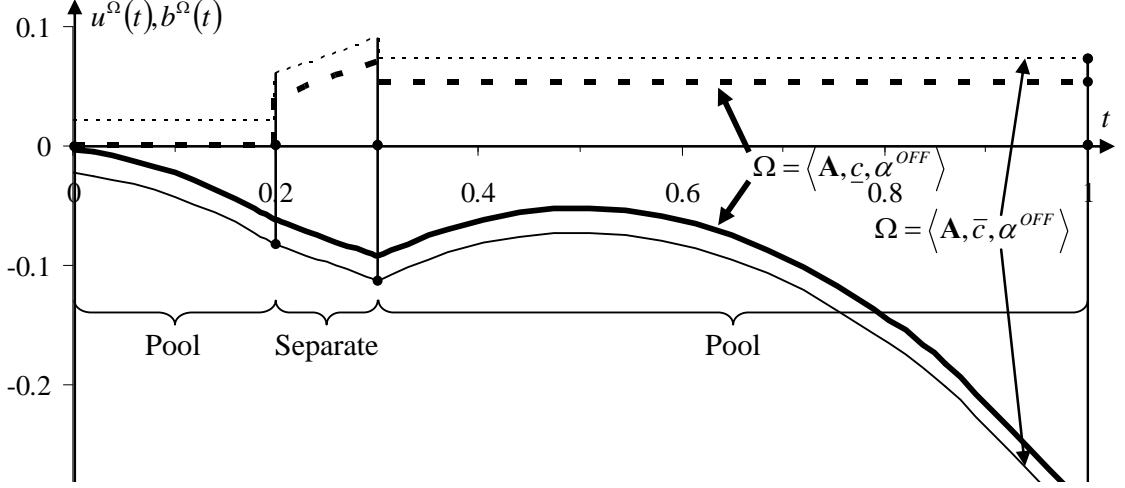


Figure 1. Two equilibrium outcomes for $\mathbf{A} = \{(0, 0.2), (0.3, 1)\}$ and $x = 0.15$.

Figure 1 also depicts a second equilibrium money burning function and concomitant Sender's utility for $c = \bar{c} = 0$ (thin lines). In this case, types $t \in (0, 0.2)$ burn $b_1^P(t) = 0.02$, types $t \in (0.2, 0.3)$ burn $b^S(t) = 0.3t$, and types $t \in (0.3, 1)$ burn $b_3^P(t) = 0.0725$. The only off equilibrium path action that supports this equilibrium is $\alpha^{OFF} = 0$. The value $c = \bar{c} = 0$ is the highest value of the parameter c for which WPBE exists: for $c > \bar{c}$ no off equilibrium path action can support this strategy profile.

Ex-ante equilibrium utilities (5) represent players' preferences over the set of all WPBE. In particular, for a given partition $\mathbf{A} \in \Pi(N)$, the Sender prefers an equilibrium with a lower value of c whereas the Receiver is indifferent. Hence, WPBE $\Omega = \langle \mathbf{A}, \underline{c}(\mathbf{A}), \alpha^{OFF} \rangle$ Pareto dominates any other WPBE $\Omega = \langle \mathbf{A}, c, \dots \rangle$ with $c > \underline{c}(\mathbf{A})$.

3. Sender's Optimal Equilibrium

We denote Sender's optimal equilibrium by Ω^S . According to Proposition 1, if $\Omega^S = \langle \mathbf{A}^S, c^S, \alpha^{OFF} \rangle$ exists it is either a solution to the following optimization problem:

$$\max_{(\mathbf{A}, c)} V^S = -c - x^2 - x + \frac{1}{6} \sum_{i=0}^{N-1} (B_i - A_i)^3$$

subject to

$$\begin{cases} \mathbf{A} \in \Pi(N) \\ c^P(A_i, B_i, c) \equiv (B_i + A_i)x - (B_i - A_i)^2 / 4 + c \geq 0 \\ (2xB_i + c)(A_{i+1} - B_i) \geq 0 \end{cases}$$

for finite N , or it is the limit thereof when the number of pooling intervals N unboundedly increases. The first inequality $c^P(A_i, B_i, c) \geq 0$ requires that the pooling types $t \in (A_i, B_i)$ burn non-negative amounts of money. The second inequality requires that if there is a positive measure of the separating types $t \in (B_i, A_{i+1})$ so that $(A_{i+1} - B_i) > 0$, each of them burns non-negative amount $b^S(t) = 2xt + c \geq 0$, which can only be binding when $t = B_i$.

It is easy to note the trade-off of the Sender. On the one hand, he benefits from having fewer and longer pooling intervals in \mathbf{A} thereby increasing the sum in the objective. On the other hand, this requires a larger value of c to maintain the inequality $c^P(A_i, B_i, c) \geq 0$ which is costly as it directly decrease the value of V^S .

We prove the existence and derive Ω^S in several steps by showing which WPBE are certainly non-optimal for the Sender. Here is the first restriction.

Lemma 1. If the measure of types that separate in WPBE Ω is positive, *i.e.*, if $B_i < A_{i+1}$, then Ω is not optimal.

According to Lemma 1, it is never in the interest of S to reveal his type perfectly by burning money. The reason is that it is too costly for the marginal types: they are strictly better-off by pooling themselves. Thus, if Ω^S does exist it must only contain pooling intervals so that $B_i = A_{i+1}$ for all $i = 1, \dots, N$. Using this result, we rewrite the optimization problem of S as follows:

$$\max_{(\mathbf{A}, c)} V^S = -c - x^2 - x + \frac{1}{6} \sum_{i=0}^{N-1} (A_{i+1} - A_i)^3 \quad \text{subject to} \quad (6)$$

$$(A_{i+1} + A_i)x - (A_{i+1} - A_i)^2 / 4 + c \geq 0, \text{ and } 0 = A_1 < \dots < A_i < A_{i+1} < \dots < A_{N+1} = 1 \quad (7)$$

The second restriction is on the number of pooling intervals N .

Lemma 2. For any $x > 0$ there exists a number $\bar{N}(x) \geq 1$ such that if WPBE Ω has $N > \bar{N}(x)$ intervals, then Ω is not optimal.

According to Lemma 2, it is never in the interest of the Sender to have infinitely many pooling intervals. The reason is very similar to that of Lemma 1. Indeed, in a partition with infinitely many intervals, some (in fact, infinitely many) of them must be of an arbitrary small length and the same logic applies: it is cheaper for the Sender to pool some of these short intervals into a bigger one. Using this result, we only need to consider finite partitions.

The next restriction is on the number of the pooling intervals where the Sender burns money.

Lemma 3. If in WPBE Ω , the Sender burns money in more than one interval, then Ω is not optimal.

According to Lemma 3, if the Sender finds money burning attractive he chooses only one pooling interval where he burns. The reason for this result is the convexity of the objective function with respect to marginal types A_i . When money is burned in two or more intervals, the number of binding restrictions $c_i^P = 0$ in the optimization problem is smaller than the number of independent variables A_i . Therefore, small deviations in at least one of them violates neither of the equilibrium conditions. However, due to the convexity of the objective function, the maximum cannot be achieved at an interior point, which implies that one additional restriction become binding. Only when all $c_i^P = 0$, except for possibly one, all A_i are fully determined by the value of c , and the maximum only occurs at the corner.

This result suggests that if Ω^S exists, its structure is very similar to the pure cheap talk equilibria in Crawford and Sobel (1982). The next restriction is on the position of the interval where S burns money in equilibrium partition.

Lemma 4. If in WPBE Ω , the Sender burns money in the interval which is not the second interval $t \in (A_2, A_3)$, then Ω is not optimal.

According to Lemma 4, if the Sender burns money in Ω^S , she only burns it in the second interval. The reason why burning money in the first interval $t \in (0, A_2)$ is suboptimal is that the objective function V^S is strictly monotone in A_2 and, hence, there is another equilibrium $\tilde{\Omega}$ that dominates Ω . On the other hand, burning money in any interval k , $k > 2$, is suboptimal because there is another equilibrium $\tilde{\Omega}$ in which S burns in the previous interval $k - 1$ and which also dominates Ω .

One important consequence of Lemma 4 is that Ω^S exists. Indeed, the choice of $A_2 \in [0, 1]$ by means of (7) uniquely defines a value of $c = c(A_2)$ by $c_1^P = c^P(0, A_2, c(A_2)) = 0$, and then, since $A_{N+1} = 1$, it defines values of A_{N-i} for $i \in \{0, \dots, N-3\}$ by $c_i^P = c^P(A_{N-i}, A_{N-i+1}, c(A_2)) = 0$. Therefore, the objective function (6) is continuous in A_2 and attains its global maximum at Ω^S .

It turns out that the structure of Ω^S is closely related to the structure of a particular equilibrium if the Sender can only use cheap talk. For every $N \in \{1, 2, 3, \dots\}$, we define x_N as

$$x_N = \frac{1}{2N(N-1)}. \quad (8)$$

Without the option to burn money, if the degree of misalignment x satisfies $x \in (x_{N+1}, x_N)$, for every value of $n \in \{1, \dots, N\}$ there exists an equilibrium in which the Sender uses n messages, see Crawford and Sobel (1982). For every x , the equilibrium with the highest

number of messages, which we denote by N^{CT} , is called the most informative equilibrium, and we denote it by Ω^{CT} :

$$N^{CT}(x) = \max\{N : x \leq x_N\}.$$

Thus, the equilibrium conditions put an upper bound on the number of messages that the Sender can use. The Sender prefers an equilibrium with the largest number of messages for a given value of x .

Intuitively, the introduction of an additional instrument to communicate – money burning – can be used to alleviate this bound. This is particularly cost-effective for those values of x for which a slight reduction in x would already have led to an increase in the number of messages in the absence of money burning. The equilibrium conditions also determine the exact positioning of the intervals of Ω^{CT} . Hence, although it may not be optimal to use money burning to increase the number of messages, it can be used to change the lengths of the intervals. This is particularly beneficial for values of x for which an additional message has just emerged, and the lengths of the intervals are unequal. These intuitions are borne out by the next proposition.

Proposition 2. For every $N \geq 4$ there exist \underline{y}_N and \bar{y}_N satisfying

$x_{N+1} < \bar{y}_{N+1} < \underline{y}_N < x_N < \bar{y}_N$, such that:

- a) If $x \in [x_N, \bar{y}_N)$ so that $N^{CT} = N - 1$, then $N^S = N^{CT} + 1 = N$ and $c_2^P > 0$, i.e., the Sender's optimal equilibrium strategy has one more interval than the most informative cheap talk equilibrium strategy, and money is burned in the second interval.
- b) If $x \in (\underline{y}_N, x_N]$ so that $N^{CT} = N$, then $N^S = N^{CT} = N$ and $c_2^P > 0$, i.e., the Sender's optimal equilibrium strategy has the same number of intervals as the most informative cheap talk equilibrium strategy, and money is burned in the second interval.
- c) Off equilibrium path beliefs in Ω^S can be taken to be $\Pr(t=0|s^{OFF})=1$ and $\alpha^{OFF}=0$.

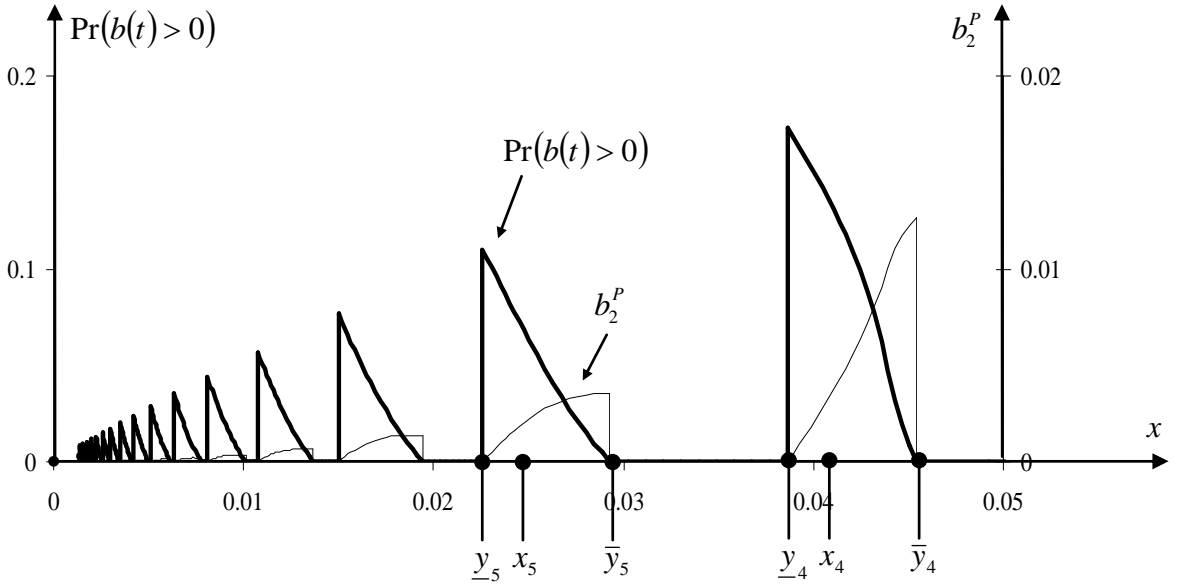


Figure 2. The amount of money burned, b_2^P , and the *ex ante* probability that it is positive, $\Pr(b(t) > 0)$, as functions of x .

Proposition 2 essentially says that the introduction of money burning raises the Sender's expected utility if x is close to x_N and if the most informative cheap talk equilibrium allows for at least three distinct messages ($N \geq 4$).

Figure 2 illustrates how, for different values of x , money burning is present or absent from the optimal equilibrium Ω^S in an alternating pattern. It shows the *ex ante* probability that money is burned $\Pr(b(t) > 0)$ (thick line, left scale), which is the length of the second interval $(A_3 - A_2)$ conditional on $b(t) > 0$, and the amount of money burned b_2^P in the second interval of Ω^S , and (thin line, right scale). Note that, for example, at $x = \bar{y}_4 = 1/22$, the length of this interval is zero, and increases with further reductions in x . At $x = \underline{y}_4 \approx 0.039$, the Sender burns money for nearly 20% of the possible values of t . The amount of money burned b_2^P is strictly positive at $x = \bar{y}_4$, and decreases with further reductions in x . This is so because equations $A_3 = A_2$, $c_1^P = c^P(0, A_2, c) = 0$ and $c_2^P = c^P(A_2, A_3, c)$ yield at $x = \bar{y}_4$:

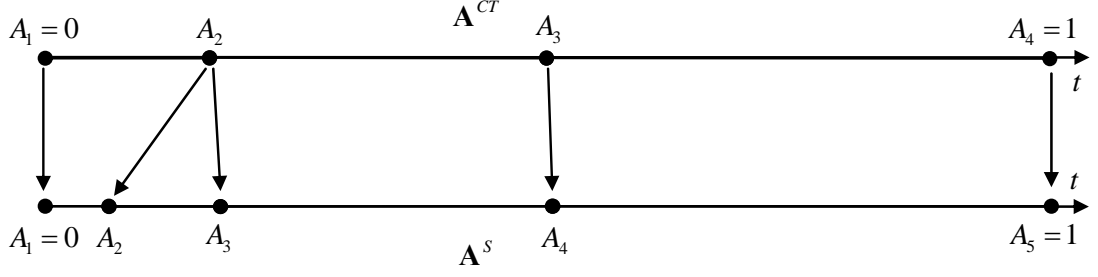


Figure 3. Equilibrium partitions \mathbf{A}^{CT} and \mathbf{A}^S for $x = 0.042$.

$$b_2^P = A_2 x + \frac{1}{4} A_2^2 > 0.$$

Money burning cannot improve upon the most informative cheap talk equilibrium for $x \geq \bar{y}_4 = 1/22$ (see the remark after the proof of the proposition in Appendix). That is, for money burning to improve the Sender's utility, the interests between Sender and Receiver should be aligned sufficiently well such that unaided cheap talk communication allows for at least three messages.

Irrespective of whether the Sender's strategy space is limited to cheap talk or also includes money burning, a decrease in x eventually gives rise to the emergence of an additional interval (message). There is, however, an interesting difference in the location of the newly emerged additional interval. If the Sender only has cheap talk at his disposition, new intervals emerge as the first interval at $t = 0$. If, in addition, the Sender also has some money to burn, new intervals appear as a second interval.

As an illustration, Figure 3 shows the equilibrium partitions, \mathbf{A}^{CT} and \mathbf{A}^S for $x = 0.042$ in the absence and the presence of money burning as a possible instrument of communication, respectively. Since $x \in (x_4, \bar{y}_4)$, the most informative cheap talk equilibrium Ω^{CT} has three intervals (with up to three digits of accuracy):

$$\mathbf{A}^{CT} = \{(0, 0.165), (0.165, 0.499), (0.499, 1)\},$$

whereas the optimal equilibrium Ω^S has four intervals:

$$\mathbf{A}^S = \{(0, 0.066), (0.066, 0.176), (0.176, 0.504), (0.504, 1)\}.$$

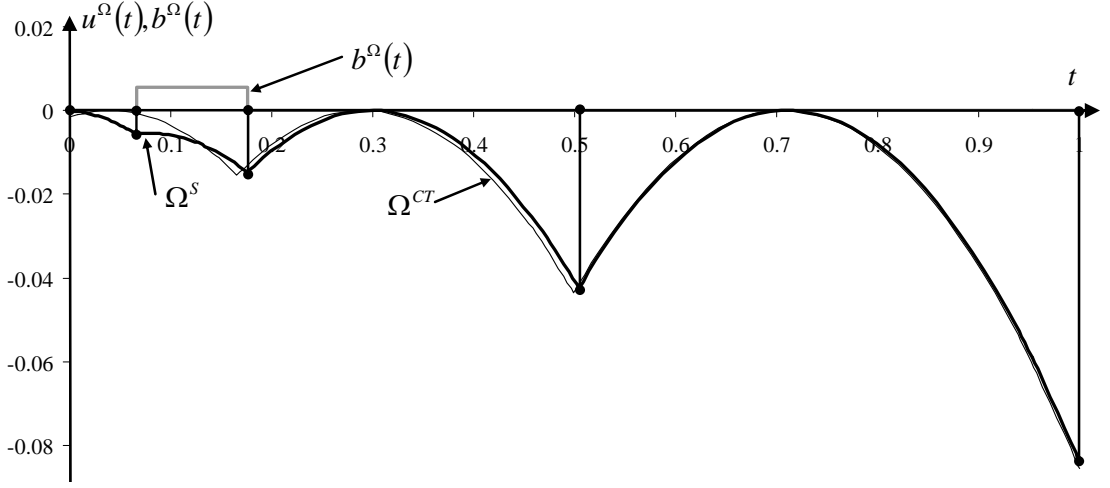


Figure 4. Equilibria Ω^{CT} and Ω^S for $x = 0.042$.

It shows how the additional interval appears at A_2 .

Figure 4 further shows, for $x = 0.042$, *ex-post* equilibrium utility functions $u^\Omega(t)$ of the Sender in Ω^{CT} and Ω^S , and the money burning function $b^\Omega(t)$: the Sender burns $b_2^P \approx 5.5 \cdot 10^{-3}$ in the second interval of \mathbf{A}^S . In this way, the expected amount of money burned $E(b) = b_2^P (A_3 - A_2) \approx 6.0 \cdot 10^{-4}$ generates a gross utility gain of $V^S + E(b) - V^{CT} \approx 7.2 \cdot 10^{-4}$. This type of equilibrium improvement is always possible as long as $x \in [x_{N+1}, \bar{y}_{N+1})$.

Figure 5, drawn for $x = 0.04$, shows how money burning is used to raise the Sender's utility by leaving the number of intervals unaffected, but changing their position. Since $x \in (\underline{y}_4, x_4)$, the most informative cheap talk equilibrium Ω^{CT} has four intervals:

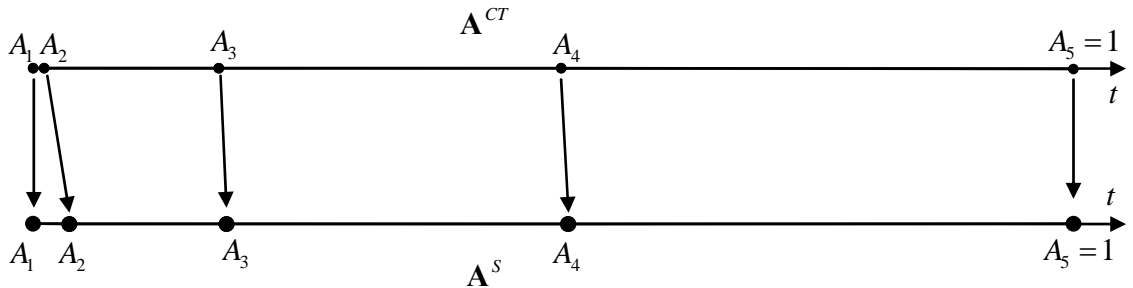


Figure 5. Equilibrium partitions \mathbf{A}^{CT} and \mathbf{A}^S for $x = 0.04$.

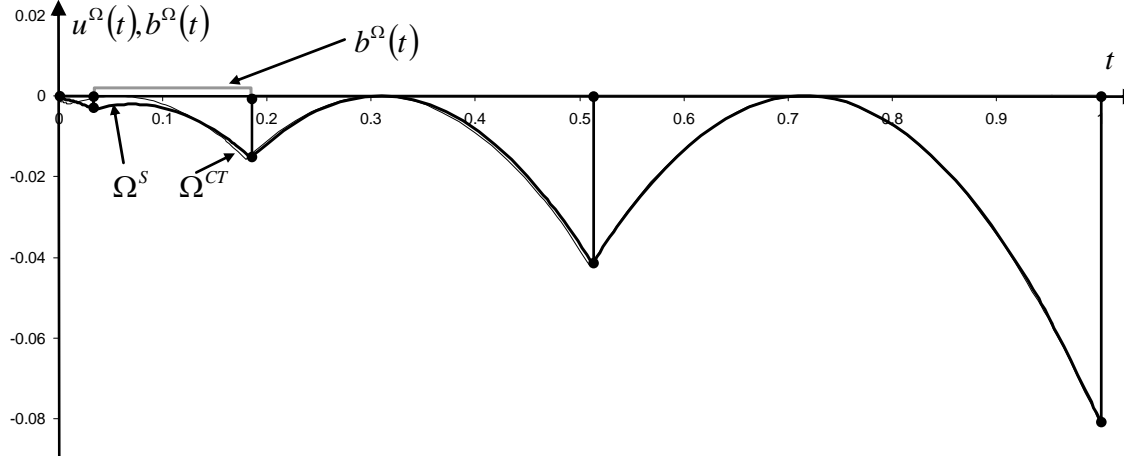


Figure 6. Equilibria Ω^{CT} and Ω^S for $x = 0.04$.

$$\mathbf{A}^{CT} = \{(0, 0.1), (0.01, 0.18), (0.18, 0.51), (0.51, 1)\},$$

and so does the optimal equilibrium Ω^S :

$$\mathbf{A}^S = \{(0, 0.035), (0.035, 0.185), (0.185, 0.512), (0.512, 1)\}.$$

Figure 6 further shows, for $x = 0.04$, *ex-post* equilibrium utility functions $u^\Omega(t)$ of the Sender in Ω^{CT} and Ω^S , and the money burning function $b^\Omega(t)$: the Sender burns $b_2^P = 2.0 \cdot 10^{-3}$ in the second interval of \mathbf{A}^S . In this way, the expected amount of money burned $E(b) = b_2^P (A_3 - A_2) \approx 3.1 \cdot 10^{-4}$ generates a gross utility gain of $V^S + E(b) - V^{CT} \approx 3.4 \cdot 10^{-4}$. This type of equilibrium improvement is always possible as long as $x \in (\underline{y}_N, x_N]$.

From the previous figures, it is not immediately clear how effective money burning is in raising the Sender's expected utility. Figure 7 plots (left scale, thick line) the relative increase in utility, $(V^S - V^{CT})/V^{CT}$ as a function of x . As one can see, the relative increase in utility is small and attains its maximum of 1.26% at $x = 0.025$. Another way of measuring the effectiveness of money burning is by comparing the expected net utility gain with the expected costs of money burning. The thin line (right scale) represents the money burning effectiveness measured by $(V^S - V^{CT})/E(b(t))$. It can be seen that for smaller values of x , money burning becomes increasingly effective in the sense that the potential utility gain from

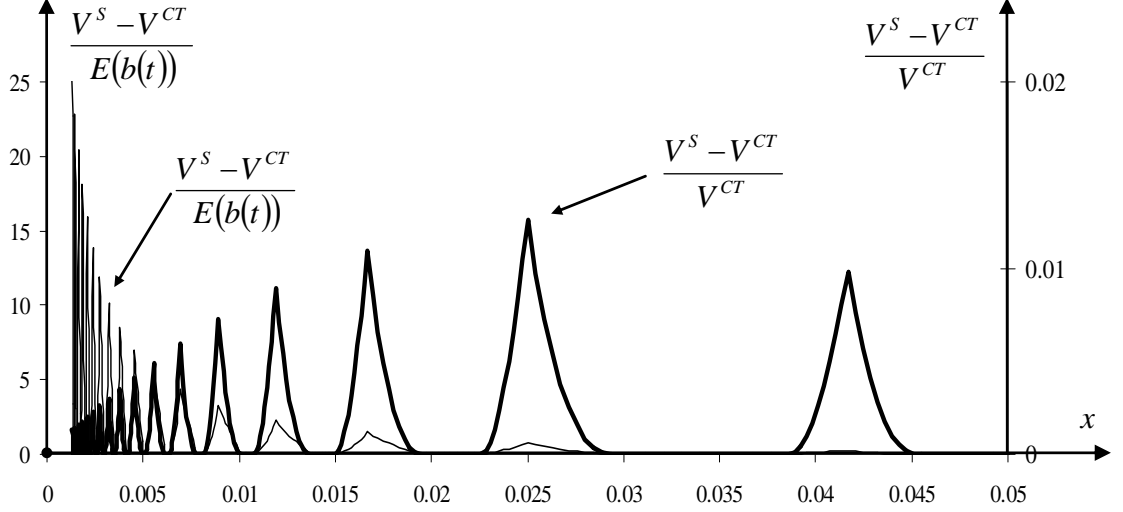


Figure 7. Relative utility gain $(V^S - V^{CT})/V^{CT}$, and the money burning effectiveness

$(V^S - V^{CT})/E(b(t))$, as functions of x .

money burning is much larger in absolute terms than the expected budget which is necessary to achieve this gain.

4. Generalizations of the Uniform-Quadratic Model

So far, we have studied in detail the uniform-quadratic model. This model is much used in applied work, and the results we obtained are therefore interesting by themselves. It is also clear that the derivations of the optimal Sender's equilibrium depend on the analytical form of the type distribution and on the Sender's utility function. Let us suppose now that the Senders' types are distributed over the type space $[0, 1]$ in accordance with a distribution function $F(t)$, which has a strictly positive and continuously differentiable density function $f(t) \equiv F'(t)$ on $[0, 1]$. Suppose also that the Sender's utility function U^S is given by:

$$U^S(a, t, s) \equiv U^S(a, t, (m, b)) = W(t, a) - b,$$

where, in addition to the assumptions in Crawford and Sobel (1982),³ its third-order partial derivatives $W_{taa}(t, a)$, and $W_{ta}(t, a)$ are continuous functions on $[0, 1]^2$.⁴

Our first result is the condition on $W(t, a)$ and $F(t)$ under which Lemma 1 generalizes, and the perfect separation of types is not optimal. We define a function $\Psi(t)$ as follows:

$$\Psi(t) \equiv \frac{2W_{taa}(t, t) + W_{ta}(t, t)}{2W_{ta}(t, t)}.$$

Proposition 3. Let a WPBE Ω have an interval of separating types $(B_i, A_{i+1}) \in \mathbf{A}(\Omega)$. If the following condition holds for some $t \in (B_i, A_{i+1})$:

$$\frac{f'}{f} - \frac{f}{F} < \Psi < \frac{f'}{f} + \frac{f}{1-F}, \quad (9)$$

then Ω is strictly suboptimal. If this condition holds for all $t \in (0, 1)$, only semi-pooling equilibria can be optimal.

Proposition 3 is a variant of Lemma 1 for the generalization of the model. It can be interpreted as follows. Suppose types $t \in (B_i, A_{i+1})$ separate in equilibrium Ω . Pooling the types from the marginal neighborhood of A_{i+1} will require, in general, that all types either above or below A_{i+1} will need to burn extra marginal amount of money. The two inequalities of condition (9) guarantee that at least one of those possibilities leads to a marginal increase of the Sender's *ex-ante* utility. When condition (9) holds for all types, any equilibrium with separating types $(B_i, A_{i+1}) \in \mathbf{A}(\Omega)$ is strictly suboptimal.

Using Proposition 3, we can extend the validity of the results obtained in Section 3 for other model specifications. It is easy to see that $\Psi(t) \equiv 0$ in the uniform-quadratic model. By

³ These are $W_{aa}(t, a) < 0$ and $W_{ta}(t, a) > 0$ on $[0, 1]^2$.

⁴ A generalization of the receiver utility function $U^R(a, t, s)$ is straightforward and does not change the current analysis provided $U^R(a, t, s)$ satisfies some mild regularity assumptions similar to those of Crawford and Sobel (1982).

our assumptions on $W(t, a)$ and $F(t)$, condition (9) will be satisfied in some open neighborhood of the uniform-quadratic model in C^3 space, *i.e.*, the space of three times continuously differentiable functions with the supremum norm $\|f\| = \sum_{n=0, \dots, 3} \sup |f^{(n)}|$. Taking into account that in the proofs of Lemma 2, Lemma 3, Lemma 4, and Proposition 2, the conclusions have been made based on strict inequalities, they can be naturally generalized.

Corollary of Proposition 3. There exists a $\delta > 0$ such that for all $W(t, a)$ and $F(t)$ satisfying $\max(\|W(t, a) + (t + x - a)^2\|, \|F(t) - t\|) < \delta$, Lemmas 1-4 hold, the optimal Sender's equilibrium exists, is unique, and is characterized by Proposition 2.

According to this result, Ω^S from Proposition 2 qualitatively remains the same for all functional forms provided the distribution $F(t)$ is close in the norm to the uniform distribution and the utility $W(t, a)$ is close to the quadratic form.

One can observe that at least one of the inequalities of condition (9) holds when t converges to zero and to one. Our last result exploits this property.

Proposition 4. There always exists a $t \in (0, 1)$ such that condition (9) holds, *i.e.*, a perfectly separating equilibrium is not optimal for the Sender.

Proposition 4 implies that the Sender never wants to use money burning to reveal his type perfectly, for any distribution of types and loss function.

5. Discussion

Above, we have derived the optimal equilibrium from the Sender's perspective. It is not hard to see that the equilibrium that maximizes the Receiver's *ex-ante* expected utility is quite different in nature. Indeed, as the Receiver does not incur the cost of money burning, the

Receiver's optimal equilibrium is characterized by complete separation through money burning by the Sender. An interesting question then becomes what equilibrium signaling strategy is best from a social welfare perspective. We focus again on the uniform-quadratic case. Let Ω^{SW} denote the efficient equilibrium, *i.e.*, the equilibrium that maximizes the social welfare function $V^{SW}(\Omega)$:

$$V^{SW}(\Omega) \equiv V^S(\Omega) + V^R(\Omega) = -c - x^2 - x + \frac{1}{12} \sum_i^N (B_i - A_i)^3.$$

A comparison with the objective function of the Sender (5) shows that the efficient equilibrium Ω^{SW} is qualitatively the same as Ω^S . All results from Section 3 continue to hold qualitatively for this new objective function, although the exact locations of marginal types in the efficient equilibrium will be different from that of the optimal Sender's equilibrium. Thus, the practical role of money burning in improving social welfare is very limited as well.

Crawford and Sobel (1982) have characterized the Sender's optimal equilibrium if the Sender can use cheap talk only. In Section 3, we have derived the optimal equilibrium when he can also use money burning. We now discuss the Sender's optimal equilibrium if his only means of communication is money burning. Consider any equilibrium signaling strategy that involves cheap talk and possibly money burning, and consider the equilibrium partition it induces. Rather than using various cheap talk messages, the Sender can burn various amounts of money. The equilibrium partition can be approximated arbitrarily closely by making the amounts of money burned vanishingly small (but still different one from the other). As a result, the loss in expected utility due to the absence of cheap talk can be made arbitrarily small. Consequently, the *optimal* equilibrium in case of money burning only does not exist; but the *supremum* of the *ex-ante* expected utility of the Sender in the absence of cheap talk equals his *ex-ante* expected utility in case of the optimal equilibrium characterized in Section 3.

Besides the reasons stated in the Introduction, there is another reason why it is interesting to derive the Sender's equilibrium signaling strategy that maximizes his *ex-ante* expected utility. In a companion paper (Karamychev and Visser, 2011) we show that an announcement by the Sender, before observing his type, that he will play the optimal equilibrium signaling strategy is credible. The notion of credibility that we use is essentially that of Farrell and Rabin (1996) – a statement is credible if it is both self-signaling and self-committing. As the Sender is the first to move in the game, he may also have an advantage in announcing which equilibrium strategy he intends to play. There is, however, no reason to believe that being the first to move in the game necessarily means that the informed player determines how his signals should be interpreted. For example, even though it is a student who takes costly actions and makes cheap talk statements about his abilities, she may be living in a society in which potential employers decide how such actions and statements are interpreted.

6. Conclusion

In this paper, we study a Sender who can signal his private information by using both cheap talk messages and money burning. Money burning significantly changes the set of possible outcomes of the signaling game. Any partition of the type space can be made an equilibrium by properly choosing amounts of money that different types burn. However, these equilibria are not equally likely to occur since the Sender has non-trivial preferences over them. In this paper, we have derived the optimal Sender's equilibrium, *i.e.*, the equilibrium that the Sender prefers the most. This equilibrium is very similar to the most informative cheap talk equilibrium, and may differ from it in two respects. First, an extra interval may be injected in between the first and the second interval, and the second, in the resulting second interval the Sender burns money to adjust the lengths of all intervals in the equilibrium partition.

We have also analyzed some asymptotical properties of this equilibrium when the misalignment between preferences of the Sender and Receiver disappears. In this case,

money burning is used in smaller amounts, less often, but with unboundedly increasing effectiveness.

References

- Austen-Smith, D. and J.S. Banks, 2000, Cheap Talk and Burned Money, *Journal of Economic Theory* 91, pp. 1-16.
- Austen-Smith, D., and J.S. Banks, 1998, Cheap Talk and Burned Money, *Center for Mathematical Studies in Economics and Management Science DP 1245*, Northwestern University
- Crawford, V., and J. Sobel, 1982, Strategic Information Transmission, *Econometrica* 50, pp. 1431-1451.
- De Haan, Th., Th. Offerman, and R. Sloof, 2011, Money Talks? An Experimental Investigation of Cheap Talk and Burned Money, *Tinbergen Institute Discussion Paper* 11-069/1.
- Farrell, J. and M. Rabin, 1996, Cheap Talk, *Journal of Economic Perspectives* 10, pp. 103-118.
- Green, J.R., and N.L. Stokey, 2007, A two-person game of information transmission, *Journal of Economic Theory* 135, pp. 90–104.
- Karamychev, V. and B. Visser, 2011, Credibility of Optimal Signaling Equilibrium, work in progress, Erasmus School of Economics and Tinbergen Institute.
- Kartik, N., 2007, A Note on Cheap Talk and Burned Money, *Journal of Economic Theory* 136, pp. 749–758.
- Spence, M., 1973, Job Market Signaling, *Quarterly Journal of Economics* 87, pp. 355–374.

Appendix

Analytical routines, long algebra transformations, and numerical computations are omitted and available upon request.

Proof of Proposition 1.

Let us fix N and a finite partition $\mathbf{A} \in \Pi(N)$. First we derive all the necessary conditions that a WPBE with partition \mathbf{A} has to satisfy. This gives us the strategy profile (σ, α) . Then,

we show that it is always possible to satisfy these conditions by choosing parameters c and α^{OFF} properly.

We begin with the Receiver. Having observed a signal s (on and off equilibrium), R believes that $\Pr(t \in \mathbf{T}(s, \sigma)) = 1$. Then, his equilibrium strategy must be:

$$\alpha(s) = \arg \max_a E(U^R(a, t) | t \in \mathbf{T}(s, \sigma)) = \arg \max_a E(-(t - a)^2 | t \in \mathbf{T}(s, \sigma)) = E(t | t \in \mathbf{T}(s, \sigma)).$$

By the definition of \mathbf{A} (3), all types $t \in (A_i, B_i)$ send messages $m^*(A_i, B_i)$ in equilibrium, burn b_i^P , and induce the same action $a(t) = E(t | t \in (A_i, B_i)) = (A_i + B_i)/2$. Hence, if type t sends a signal $\sigma(y)$ as if his type were $y \in (A_i, B_i)$, he gets the following pooling utility:

$$u^P(t, y) \equiv U^S(a(y), t, \sigma(y)) = -(t + x - (A_i + B_i)/2)^2 - b_i^P, \text{ for } y \in (A_i, B_i). \quad (10)$$

Similarly, all types $t \in (B_i, A_{i+1})$ send individual messages $m^*(t)$, burn $b^S(t)$, and induce actions $a(t) = E(t | t) = t$. Hence, if type t sends a signal $\sigma(y)$ as if his type were $y \in (B_i, A_{i+1})$, he gets the following utility:

$$u^S(t, y) \equiv U^S(a(y), t, \sigma(y)) = -(t + x - y)^2 - b^S(y).$$

In equilibrium, each type $t \in (B_i, A_{i+1})$ must find it optimal to signal his true type. Hence, it maximizes $u(t, y)$ with respect to y , and the truth telling condition requires that the maximum occurs at $y = t$. This yields the first-order condition $0 = 2x - (b^S)'(t)$ which solution is $b^S(t) = 2xt + b_i^S$, where b_i^S is an unknown constant. Thus, if type t sends a signal $\sigma(y)$ with $y \in (A_i, B_i)$, its separating utility is

$$u^S(t, y) = U^S(a(y), t, \sigma(y)) = -(t + x - y)^2 - 2xy - b_i^S, \text{ for } y \in (B_i, A_{i+1}). \quad (11)$$

Combining (10) and (11) we write utility of type t when its signal is $\sigma(y)$ as follows:

$$u(t, y) = \begin{cases} -(t + x - (A_i + B_i)/2)^2 - b_i^P & , \text{ if } y \in (A_i, B_i) \\ -(t + x - y)^2 - 2xy - b_i^S & , \text{ if } y \in (B_i, A_{i+1}) \end{cases} \quad (12)$$

Each type $t \in [0, 1]$ maximizes $u(t, y)$ with respect to y . Hence, the equilibrium utility $u^\Omega(t)$ can be written as follows:

$$u^\Omega(t) = \max_{y \in [0, 1]} u(t, y) = u(t, t).$$

Thus, $u^\Omega(t)$ is the upper envelope of $u(t, y)$ and, therefore, is continuous in t at all intervals' boundaries $t = A_i$ and $t = B_i$. This allows us to relate b_i^P , b_i^S , A_i , and B_i in the indifference equations $u^P(A_i, A_i) = u^S(A_i, A_i)$ and $u^P(B_i, B_i) = u^S(B_i, B_i)$, which leads to the following relation:

$$b_{i-1}^S = b_i^S = (B_i - A_i)^2 / 4 - (B_i + A_i)x + b_i^P.$$

This implies that all b_i^S are identical, and we denote its common value by c : $b_i^S = c$. Using this notation and (4), the above relation reads as $b_i^P = c^P(A_i, B_i, c)$, and the equilibrium utility $u^\Omega(t)$ can be written as

$$u^\Omega(t) = \begin{cases} -x^2 - 2xt - c + (B_i - t)(t - A_i) & , \text{ if } t \in (A_i, B_i) \\ -x^2 - 2xt - c & , \text{ if } t \in (B_i, A_{i+1}) \end{cases}$$

This proves parts (a), (b), (c), and (e) of the proposition, except for the off equilibrium action α^{OFF} . The *ex-ante* expected utility of S and R are as follows:

$$V^S(\Omega) = \int_0^1 u^\Omega dt = \int_0^{A_1} u^\Omega dt + \sum_{i=1}^N \left(\int_{A_i}^{B_i} u^\Omega dt + \int_{B_i}^{A_{i+1}} u^\Omega dt \right) = -c - x^2 - x + \frac{1}{6} \sum_{i=1}^N (B_i - A_i)^3,$$

$$V^R(\Omega) = - \int_0^1 (t - a(t))^2 dt = - \sum_{i=1}^N \int_{A_i}^{B_i} \left(t - \frac{1}{2}(A_i + B_i) \right)^2 dt = - \frac{1}{12} \sum_{i=1}^N (B_i - A_i)^3.$$

This proves part (d) of the proposition.

In equilibrium, S can only burn non-negative amounts of money. By taking $c = 1/4$ we can ensure that $b_i^P > 0$ and $b^S(t) > 0$. Then, we start reducing the value of c until either $b_i^P = 0$ for some i , or $b^S(B_i) = 0$ for some i for which $B_i < A_{i+1}$. This can always be done as c is an additive constant in b_i^P and $b^S(t)$. This defines $\underline{c}(\mathbf{A})$:

$$\underline{c}(\mathbf{A}) = \min \left\{ c : \left(c^P(A_i, B_i, c) \geq 0 \text{ for } \forall i \right) \text{ AND } \left((A_{i+1} - B_i)(2xB_i + c) \geq 0 \text{ for } \forall i \right) \right\}.$$

For any $c < \underline{c}(\mathbf{A})$, some types would have need to burn negative amount of money, which is not feasible. At $c = \underline{c}(\mathbf{A})$, there is a type t_0 that burns no money, and for $c > \underline{c}(\mathbf{A})$ all types burn. By the construction of $\underline{c}(\mathbf{A})$, the equilibrium $\Omega = \langle \mathbf{A}, c, \alpha^{OFF} \rangle$ with $c = \underline{c}(\mathbf{A})$ and $\alpha^{OFF} = t_0$ always exists. Finally, we define $\bar{c}(\mathbf{A})$ as the highest value of c for which we can still find $\alpha^{OFF} \in [0, 1]$ that support WPBE $\Omega = \langle \mathbf{A}, c, \alpha^{OFF} \rangle$. By construction, $\bar{c}(\mathbf{A}) \geq \underline{c}(\mathbf{A})$.

This completes the proof. Q.E.D.

Proof of Lemma 1.

Let us fix a WPBE Ω with partition $\mathbf{A}(\Omega)$ in which $A_{i+1} > B_i$ for some i so that all types $t \in (B_i, A_{i+1})$ burn $b^S(t) = 2xt + c \geq 0$ in equilibrium. We will show that there is another equilibrium $\tilde{\Omega} \neq \Omega$ (we use the tilde sign ‘ \sim ’ to denote all its variables) that generates a higher level of utility, *i.e.*, that $V^S(\tilde{\Omega}) > V^S(\Omega)$, so that Ω is non-optimal.

Let us take an $\varepsilon \in (0, \min(4x, A_{i+1} - B_i))$, and consider the equilibrium with partition $\mathbf{A}(\tilde{\Omega}) = \mathbf{A}(\Omega) \cup \{(B_i, B_i + \varepsilon)\}$, which only differs from $\mathbf{A}(\Omega)$ in that types $t \in (B_i, B_i + \varepsilon)$ pool together. In Ω , $b^S(B_i) = 2xB_i + c \geq 0$ so that types B_i and $B_i + \varepsilon$ get utilities

$$u^\Omega(B_i) = -x^2 - 2xB_i - c \text{ and } u^\Omega(B_i + \varepsilon) = -x^2 - 2x(B_i + \varepsilon) - c.$$

In $\tilde{\Omega}$, types $t \in (B_i, B_i + \varepsilon)$ burn $\tilde{b}^S(t) = c^P(B_i, B_i + \varepsilon, c)$:

$$c^P(B_i, B_i + \varepsilon, c) = (2B_i + \varepsilon)x - \varepsilon^2 / 4 + c = 2xB_i + c + (4x - \varepsilon)\varepsilon / 4 > 0,$$

so that the constraint $c^P(B_i, B_i + \varepsilon, c) \geq 0$ is not violated in $\tilde{\Omega}$. Types $t \in (B_i, B_i + \varepsilon)$ get utilities:

$$u^{\tilde{\Omega}}(t) = -x^2 - 2xt - c + t((B_i + \varepsilon) + B_i) - (B_i + \varepsilon)B_i - t^2.$$

This implies that the marginal types B_i and $B_i + \varepsilon$ get

$$u^{\tilde{\Omega}}(B_i) = -x^2 - 2xB_i - c = u^{\Omega}(B_i) \text{ and } u^{\tilde{\Omega}}(B_i + \varepsilon) = -x^2 - 2x(B_i + \varepsilon) - c = u^{\Omega}(B_i + \varepsilon),$$

i.e., the same utilities as they get in Ω . Thus, all equilibrium conditions for $\tilde{\Omega}$ are satisfied.

However, in WPBE $\tilde{\Omega}$, the objective function $V^S(\tilde{\Omega})$ given by (5) has one extra term, namely $\varepsilon^3/6$ that corresponds to an extra pooling interval $(B_i, B_i + \varepsilon)$, as compared to WPBE Ω .

Thus, $V^S(\tilde{\Omega}) = V^S(\Omega) + \varepsilon^3/6$ and, therefore, Ω is not optimal

Q.E.D.

Proof of Lemma 2.

Let us take any WPBE Ω , which partition $\mathbf{A}(\Omega) = \{(A_i, B_i) | i = 1, \dots, \infty\}$ has infinitely countable number of intervals (we use the notation B_i for such a partition because the existence of A_{i-1} or A_{i+1} is not guaranteed for every A_i as the index set \mathbf{I} can be any ordered set). The set of marginal types $\{A_i\}_{i \in \mathbf{I}}$ has at least one accumulation point, and we take any of them and denote by \hat{A} . It follows that for any $\varepsilon > 0$, there are infinitely many marginal types either in its right neighborhood $(\hat{A}, \hat{A} + \varepsilon)$ or left neighborhood $(\hat{A} - \varepsilon, \hat{A})$. We consider the first possibility; the other case is treated similarly.

Let us take an $\varepsilon \in (0, 4x)$ and choose two marginal types A_{n_1} and A_{n_2} such that $n_1 < n_2$ and $\hat{A} < A_{n_1} < A_{n_2} < \hat{A} + \varepsilon$. In Ω , marginal type A_i burns $b^S(A_i) = 2xA_i + c$. By continuity of $u^\Omega(t)$, type \hat{A} gets utility

$$u^\Omega(\hat{A}) = \lim_{k \rightarrow \infty} u^\Omega(A_{i_k}) = -x^2 - 2\hat{A}x - c,$$

where the limit is taken over any decreasing subsequence of $\{A_{i_k}\}$ that converges to \hat{A} . By construction, type \hat{A} gets the same utility if he perfectly reveals his type, in which case he burns $b^S(\hat{A}) = 2x\hat{A} + c \geq 0$. In Ω , S gets the following utility $V^S \equiv V^S(\Omega)$:

$$V^S = -c - x^2 - x + \frac{1}{6} \left(\sum_{A_i < \hat{A}} (B_i - A_i)^3 + \sum_{\substack{A_i \geq \hat{A} \\ A_i < A_{n_1}}} (B_i - A_i)^3 + \sum_{\substack{A_i \geq A_{n_1} \\ A_i < A_{n_2}}} (B_i - A_i)^3 + \sum_{A_i \geq A_{n_2}} (B_{i+1} - A_i)^3 \right).$$

Let us now consider WPBE $\tilde{\Omega}$ such that in the underlying partition $\mathbf{A}(\tilde{\Omega})$, all types $t \in (\hat{A}, A_{n_2})$ pool together and all other types follow the same strategies as in Ω . In $\tilde{\Omega}$, S gets:

$$\tilde{V}^S \equiv V^S(\tilde{\Omega}) = -c - x^2 - x + \frac{1}{6} \left(\sum_{A_i < \hat{A}} (B_i - A_i)^3 + (A_{n_2} - \hat{A})^3 + \sum_{A_i \geq A_{n_2}} (B_i - A_i)^3 \right).$$

This implies:

$$\begin{aligned} \tilde{V}^S - V^S &= \frac{1}{6} \left((A_{n_2} - \hat{A})^3 - \sum_{\hat{A} \leq A_i < A_{n_1}} (B_i - A_i)^3 - \sum_{A_{n_1} \leq A_i < A_{n_2}} (B_i - A_i)^3 \right) \\ &\geq \frac{1}{6} \left((A_{n_2} - \hat{A})^3 - (A_{n_1} - \hat{A})^3 - (A_{n_2} - A_{n_1})^3 \right) = \frac{1}{2} (A_{n_2} - A_{n_1}) (A_{n_1} - \hat{A}) (A_{n_2} - \hat{A}) > 0 \end{aligned}$$

In $\tilde{\Omega}$, types $t \in (\hat{A}, A_{n_2})$ burn:

$$\begin{aligned} \tilde{b}^S(t) &= c^P(\hat{A}, A_{n_2}, c) = (A_{n_2} + \hat{A})x - \frac{1}{4}(A_{n_2} - \hat{A})^2 + c \\ &= (2x\hat{A} + c) + \frac{1}{4}(A_{n_2} - \hat{A})(4x - (A_{n_2} - \hat{A})) = b^S(\hat{A}) + \frac{1}{4}(A_{n_2} - \hat{A})(4x - (A_{n_2} - \hat{A})) \\ &> b^S(\hat{A}) + \frac{1}{4}(A_{n_2} - \hat{A})(4x - \varepsilon) > b^S(\hat{A}) > 0 \end{aligned}$$

Thus, all equilibrium conditions for $\tilde{\Omega}$ are satisfied, and Ω is suboptimal due to $V^S(\tilde{\Omega}) - V^S(\Omega) > 0$. This ends the proof. *Q.E.D.*

Proof of Lemma 3.

Let us fix a WPBE Ω in which the Sender burns money in some two intervals and does not burn in between them. We denote the index of the second such interval by k , $k = 2, \dots, N$, and the number of intervals in between by n , $n = 0, \dots, k-2$. In other words, we assume that

$$c_k^P > 0, c_{k-n-1}^P > 0 \text{ and } c_i^P = 0 \text{ for } i = (k-n), \dots, (k-1).$$

We will show that there is another equilibrium $\tilde{\Omega}$ such that $V^S(\tilde{\Omega}) - V^S(\Omega) > 0$. In particular, we consider a family of equilibria which have the same value of the parameter c , and the same marginal types in the set $[0, A_{k-n-1}] \cup [A_{k+1}, 1]$. It is easy to see that the choice of A_k uniquely defines A_i for $i = (k-n), \dots, (k-1)$ through $c_i^P = c^P(A_i, A_{i+1}, c) = 0$ as follows:

$$A_i = A_{i+1} - 2\left(\sqrt{x^2 + 2xA_{i+1} + c} - x\right).$$

By induction on i , this recursive equation implies:⁵

$$A_{k-j} = A_k - 2j(s(A_k, c) - jx) \text{ for } j=1, \dots, n, \quad (13)$$

where

$$s(A, c) \equiv \sqrt{x^2 + 2xA + c}.$$

Therefore, the lengths of the intervals, which we denote by

$$D_i \equiv A_{i+1} - A_i,$$

can be written as follows:

$$D_{k-j} = A_{k-j+1} - A_{k-j} = 2(s(A_k, c) - (2j-1)x) > 0 \text{ for } j=1, \dots, n.$$

The constraints $c^P(A_i, A_{i+1}, c) = 0$ imply that:

$$D_{k-j}^3 = (A_{k-j+1} - A_{k-j})^3 = 4x(A_{k-j+1}^2 - A_{k-j}^2) + 4c(A_{k-j+1} - A_{k-j}) \text{ for } j=1, \dots, n. \quad (14)$$

In this notation, V^S can be written as the following function of A_k :

$$V^S \equiv \hat{V}^S(A_k) = -c - x^2 - x + \frac{1}{6} \left(\sum_{i=1}^{k-n-2} D_i^3 + D_{k-n-1}^3 + \sum_{j=1}^n D_{k-j}^3 + D_k^3 + \sum_{i=k+1}^N D_i^3 \right).$$

It can be shown that $\hat{V}^S(A_k)$ is convex so that it does not have an interior maximum, and its maximum necessarily violates one of the restrictions $D_{k-n-1} > 0$, $D_k > 0$, $c_k^P > 0$, or

$c_{k-n-1}^P > 0$. Consequently, in any open neighborhood of A_k , there exists an \tilde{A}_k so that the corresponding WPBE $\tilde{\Omega}$ is such that $V^S(\tilde{\Omega}) - V^S(\Omega) > 0$, which ends the proof. *Q.E.D.*

Proof of Lemma 4.

It is easy to see that if $N = 1$ so that $A(\Omega) = \{(0, 1)\}$, and $c_1^P > 0$, Ω is not optimal. Indeed, by reducing the value of c , we get the same completely pooling WPBE in which less money

⁵ The necessary condition for the induction is that $s(A, c) - 2nx > 0$. When it fails, A_{k-j} does not exist for some $j=1, \dots, n$. Hence, this condition is necessarily satisfied by the assumption that $c_i^P = 0$ for $i = (k-n), \dots, (k-1)$.

is burned. In the rest of the proof, we assume that $N \geq 2$. The proof consists of two parts.

First, we show that Ω is not optimal if $c_1^P > 0$. Second, we show that neither is Ω optimal if $c_i^P > 0$ for $i \in \{3, \dots, N\}$.

Part 1. Let $c_1^P > 0$ in Ω . Then, in accordance with Lemma 3, $c_i^P = 0$ for all $i \in \{2, \dots, N\}$.

Constraints $c_i^P = 0$ define c and A_i as functions of A_N through (13):

$$\begin{aligned} c(A_N) &= (1 - A_N)^2 / 4 - (1 + A_N)x, \\ A_i(A_N) &= 1 - (N + 1 - i)((1 - A_N) - 2(N - i)x) \text{ for } i \in \{2, \dots, N\}. \end{aligned} \quad (15)$$

Then, (14) implies that

$$V^S(A_N) = -c - x^2 - x + \frac{1}{6} \left(D_1^3 + \sum_{i=2}^N D_i^3 \right) = -c - x^2 - x + \frac{1}{6} (A_2^3 + 4x(1 - A_2^2) + 4c(1 - A_2)),$$

and its derivative

$$\frac{dV^S}{dA_N}(A_N) = \frac{(N-1)}{6} \left(3A_2^2 - 8xA_2 - (1 - A_N)^2 + 4(1 + A_N)x - \frac{(2(1 - A_2) - 3)}{(N-1)}((1 - A_N) + 2x) \right).$$

It can be shown that $dV^S/dA_N > 0$. Hence, $V^S(A_N)$ strictly increases in A_N and, therefore,

Ω is not optimal. This ends the proof of part 1.

Part 2. Let $c_{k+1}^P > 0$ for some $k \in \{2, \dots, N-1\}$ in Ω . Then, in accordance with Lemma 3,

$c_i^P = 0$ for $i \in \{1, \dots, k\}$, which defines c and A_i for $i \in \{2, \dots, k+1\}$ as functions of A_2 :

$$c = A_2^2 / 4 - xA_2, \text{ and } A_i = (i-1)(2x(i-2) + A_2).$$

Next, we formally define \hat{A}_{k+2} by $c^P(A_{k+1}, \hat{A}_{k+2}, c) = 0$:

$$\hat{A}_{k+2} \equiv (k+1)(2kx + A_2).$$

We will show that $A_{k+2} < \hat{A}_{k+2}$. To this end, we observe that the function $c^P(A, B, c)$ is the second-degree concave polynomial w.r.t. B . In addition,

$$c^P(A_{k+1}, A_{k+1}, c) = 2xA_{k+1} + c = 4x^2k(k-1) + (2k-1)xA_2 + A_2^2 / 4 > 0,$$

since $k \geq 2$, and also $c^P(A_{k+1}, \hat{A}_{k+2}, c) = 0$ by construction. Therefore, the assumed in part 2 inequality $c_{k+1}^P = c^P(A_{k+1}, A_{k+2}, c) > 0$ can only be satisfied for $A_{k+2} < \hat{A}_{k+2}$.

The contribution of intervals (A_k, A_{k+1}) and (A_{k+1}, A_{k+2}) into the objective $V^S(\Omega)$ is $(A_{k+1} - A_k)^3 / 6 + (A_{k+2} - A_{k+1})^3 / 6$. We will show that there is another equilibrium $\tilde{\Omega}$ which generates a higher level of V^S . In particular, we consider the equilibrium with the same marginal types as in Ω except for \tilde{A}_{k+1} . In $\tilde{\Omega}$, we take \tilde{A}_{k+1} by $c^P(\tilde{A}_{k+1}, A_{k+2}, c) = 0$:

$$\tilde{A}_{k+1} \equiv A_{k+2} - 2\left(\sqrt{x^2 + 2xA_{k+2} + c} - x\right).$$

The following properties of \tilde{A}_{k+1} can be shown. First, that $\tilde{A}_{k+1} \in (A_k, A_{k+1})$ and $\tilde{c}_k^P = c^P(A_k, \tilde{A}_{k+1}, c) > 0$ so that $\tilde{\Omega}$ is based on a legitimate partition. Second, the contribution of intervals (A_k, \tilde{A}_{k+1}) and $(\tilde{A}_{k+1}, A_{k+2})$ into the objective \tilde{V}^S is $(\tilde{A}_{k+1} - A_k)^3 / 6 + (A_{k+2} - \tilde{A}_{k+1})^3 / 6$. Hence,

$$\begin{aligned} \tilde{V}^S - V^S &= \left((\tilde{A}_{k+1} - A_k)^3 + (A_{k+2} - \tilde{A}_{k+1})^3 - (A_{k+1} - A_k)^3 - (A_{k+2} - A_{k+1})^3 \right) / 6 \\ &= (A_{k+1} - \tilde{A}_{k+1})(A_{k+2} - A_k)((A_{k+2} - \tilde{A}_{k+1}) - (A_{k+1} - A_k)) / 2 \\ &= (A_{k+1} - \tilde{A}_{k+1})(A_{k+2} - A_k) \left(\sqrt{((2k-1)x + A_2/2)^2 + 2xv - ((2k-1)x + A_2/2)} \right) \end{aligned}$$

so that $\tilde{V}^S - V^S > 0$. Thus, Ω is suboptimal, which ends the proof of part 2. *Q.E.D.*

Proof of Proposition 2.

In what follows, we assume $N^{CT} \geq 3$ and define

$$\hat{x}_N \equiv 1/(2N^2),$$

so that $\hat{x}_N \in (x_{N+1}, x_N)$. For given x , $x \in [x_{N^{CT}+1}, x_{N^{CT}})$, the most informative cheap talk equilibrium Ω^{CT} has underlying partition $\mathbf{A} = \{(A_i, A_{i+1}) | i = 1, \dots, N^{CT}\}$ with N^{CT} intervals of positive lengths.

Suppose that WPBE Ω is the Sender's optimal equilibrium, contains N intervals, and its first marginal type is $A_2 \in (0,1]$. Using $c_1^P = c_i^P = 0$ for $i = 3, \dots, N$, we write the parameter c and the other marginal types A_i as follows:

$$c(A_2) = \frac{1}{4}A_2^2 - xA_2, \text{ and } A_i(A_2) = 1 - 2(N - i + 1)(s(A_2) - (N - i + 1)x),$$

where

$$s(A_2) = \sqrt{x^2 + 2x + c} = \frac{1}{2}\sqrt{(2x - A_2)^2 + 8x}.$$

The last two necessary condition for Ω are $A_3(A_2) > A_2$ and $c_2^P = c^P(A_2, A_3, c) \geq 0$. It can be shown that these two inequalities lead to the following restriction on A_2 :

$$\bar{A}_2 \equiv 2(N - 2)(x_{N-1} - x) > A_2 \geq 2(N - 1)(x_N - x) \equiv \underline{A}_2. \quad (16)$$

Since $x < x_{N^{CT}}$, $(x_{N-1} - x) > A_2 > 0$ requires $x_{N-1} \geq x_{N^{CT}}$ and, hence, $N \leq N^{CT} + 1$. Thus, optimal equilibrium Ω^S might only have $N^{CT} + 1$ intervals at most, and we construct it as follows. First, we fix any $N = 1, \dots, N^{CT} + 1$. Then, we consider any $A_2 \in (0,1]$ such that it satisfies (16). Third, we choose $A_2 \in (0,1]$ such that $V^S(A_2)$ is maximized:

$$\begin{aligned} V^S(A_2) &= -c - x^2 - x + \frac{1}{6} \sum_{i=1}^N (A_{i+1} - A_i)^3 \\ &= -x^2 - x + \frac{1}{6} \left(A_2^3 + (A_3 - A_2)^3 + 4x(1 - A_3^2) - 2(1 + 2A_3) \left(\frac{1}{4}A_2 - x \right) A_2 \right). \end{aligned}$$

If it happens at the corner $A_2 = \underline{A}_2$ (the other corner $A_2 = 1$ corresponds to the pooling WPBE and never occurs when $N^{CT} \geq 2$), the resulting WPBE is the cheap talk equilibrium with N interval. Obviously, $N = N^{CT}$ maximizes V^S across all N . If, on the other hand, there is an interior maximum $A_2^S \in (\underline{A}_2, \bar{A}_2)$, it implies that $V^S(A_2^S) > V^S(\underline{A}_2)$. In other words, there is a WPBE Ω^S that yields a higher utility to the sender than the cheap talk equilibrium with N intervals. In this case, the derivative of $V^S(A_2)$, which we denote by $G(A_2)$, vanishes:

$$G \equiv \frac{dV^s}{dA_2} = \frac{1}{6} \left(A_3(4(A_2 + x) - 3A_3) - A_2 + 2x + (3(A_3 - A_2)^2 - A_2^2 - 4x(2A_3 - A_2)) \frac{dA_3}{dA_2} \right),$$

where

$$\frac{dA_3}{dA_2}(A_2) = \frac{(N-2)(2x - A_2)}{2s}.$$

Since $x \in [x_{N^{CT}+1}, x_{N^{CT}})$, there can be two cases, $x \in [\hat{x}_{N^{CT}}, x_{N^{CT}})$ and $x \in [x_{N^{CT}+1}, \hat{x}_{N^{CT}})$, which we consider separately.

Part 1. Let $x \in [\hat{x}_{N^{CT}}, x_{N^{CT}})$. Suppose, first, $N = N^{CT} + 1 \geq 4$. In this case, $x \in [\hat{x}_{N-1}, x_{N-1})$, $x > \hat{x}_{N-1} > x_N$ and, therefore, $\underline{A}_2 < 0$ so that A_2 must satisfy $A_2 \in [0, \bar{A}_2)$. It can be shown that $G(A_2) > 0$ for all $A_2 \in [0, \bar{A}_2)$. Hence, $V^s(A_2)$ monotonically increases and has no maximum.

Suppose now that $N = N^{CT} \geq 3$. In this case, $x \in [\hat{x}_N, x_N)$ and A_2 must satisfy $A_2 \in [\underline{A}_2, \bar{A}_2)$. It can be shown that $G(\underline{A}_2) > 0 > G(\bar{A}_2)$ at $x = x_N$ and $G(\underline{A}_2) < 0$ at $x = \hat{x}_N$. By continuity, this implies that there exists a $\underline{y}_N \in (\hat{x}_N, x_N)$ such that $G(\underline{A}_2) > 0 > G(\bar{A}_2)$ for all $x \in (\underline{y}_N, x_N]$. This implies, in turn, that there exists $A_2^S \in (\underline{A}_2, \bar{A}_2)$ that is a local maximum of $V^s(A_2)$, and at which $c_2^P > 0$. Next, it can be shown $c(A_2) > 0$ implies $G(A_2) < 0$ for $x \in [\hat{x}_N, x_N)$. Thus, when $A_2^S \in (\underline{A}_2, \bar{A}_2)$, it is necessarily the case that $c(A_2) \leq 0$. When the maximum is at the corner $A_2 = \underline{A}_2$, it is the cheap talk most informative equilibrium with $N = N^{CT}$ intervals with

$$c(\underline{A}_2) = ((N-1)x_N - (N+1)x)(N-1)(x_N - x) = (N^2 - 1)(x_{N+1} - x)(x_N - x) < 0.$$

Finally, suppose that $N \leq N^{CT} - 1$. In this case, $x < x_{N+1}$. It can be shown that $G(A_2) < 0$ for all $A_2 \in [\underline{A}_2, \bar{A}_2)$. Hence, $V^s(A_2)$ monotonically decrease and achieves its maximum at $A_2 = \underline{A}_2$, the cheap talk equilibrium with $N < N^{CT}$ intervals.

Summarizing Part 1, if $x \in (\underline{y}_{N^{CT}}, x_{N^{CT}}]$, the maximum is in the interior, $A_2^S \in (\underline{A}_2, \bar{A}_2)$, the number of intervals is $N = N^{CT}$, and $c(A_2) \leq 0$. If, on the other hand, $x \in [\hat{x}_{N^{CT}}, \underline{y}_{N^{CT}})$ the maximum is at the corner, $A_2 = \underline{A}_2$, it is the cheap talk equilibrium with $N = N^{CT}$ intervals and $c(A_2) < 0$.

Part 2. Let $x \in [x_{N^{CT}+1}, \hat{x}_{N^{CT}})$. Suppose, first, $N = N^{CT} + 1 \geq 4$. In this case, $x \in [x_N, \hat{x}_{N-1})$, $x \geq x_N$ and, therefore, $\underline{A}_2 \leq 0$ so that A_2 must satisfy $A_2 \in [0, \bar{A}_2)$. It can be shown that $G(0) > 0 > G(\bar{A}_2)$ at $x = x_N$ and $G(\bar{A}_2) > 0$ at $x = \hat{x}_{N-1}$. By continuity, this implies that there exists a $\bar{y}_N \in (x_N, \hat{x}_{N-1})$ such that $G(0) > 0 > G(\bar{A}_2)$ for all $x \in [x_N, \bar{y}_N)$. This implies, in turn, that there exists $A_2 \in (0, \bar{A}_2)$ that is a local maximum of $V^S(A_2)$, and at which $c_2^P > 0$. Finally, since $A_2 \in (0, \bar{A}_2)$, $\bar{A}_2 - 4x = 2N(x_N - x) \leq 0$ implies $c(A_2) \leq 0$. The maximum at the corner $A_2 = \bar{A}_2$ is not feasible.

Suppose now that $N \leq N^{CT}$. In this case, $x < \hat{x}_N$. It can be shown that $G(A_2) < 0$ for all $A_2 \in [0, \bar{A}_2)$. Hence, $V^S(A_2)$ monotonically decrease and achieves its maximum at $A_2 = \underline{A}_2$, the cheap talk equilibrium with $N \leq N^{CT}$ intervals. Consequently, optimal equilibrium has $N = N^{CT}$, and $c(A_2) < 0$, as is shown in Part 1.

Summarizing Part 2, if $x \in [x_{N^{CT}+1}, \bar{y}_{N^{CT}+1})$, the maximum is in the interior, $A_2^S \in (\underline{A}_2, \bar{A}_2)$, the number of intervals is $N = N^{CT} + 1$, and $c(A_2) \leq 0$. If, on the other hand, $x \in (\bar{y}_{N^{CT}+1}, \hat{x}_{N^{CT}})$ the maximum is at the corner, $A_2 = \underline{A}_2$, it is the cheap talk equilibrium with $N = N^{CT}$ intervals and $c(A_2) < 0$.

Combining the above cases yield the proof of parts (a) and (b) of the proposition, and that $c(A_2) \leq 0$ in the optimal equilibrium. In order to prove part (c) we set $\alpha^{OFF} = 0$ and note that

$c(A_2) \leq 0$ implies that type $t = 0$ prefers sending his equilibrium message and getting utility $u^\Omega(t) = -(x - A_2/2)^2$ to sending any off equilibrium message and getting $u^{OFF}(t) = -x^2$:

$$u^{OFF}(t) - u^\Omega(t) = \frac{1}{4} A_2^2 - x A_2 = c(A_2) \leq 0.$$

All other types $t > 0$, by construction of WPBE, prefer off equilibrium messages even less. Hence, off equilibrium path beliefs with the resulting action $\alpha^{OFF} = 0$ support the optimal equilibrium of the Sender as Weak Perfect. This proves part (c) of the proposition. *Q.E.D.*

Remark. Proposition 2 provides the sufficient condition for $c_2^P > 0$. In order to show that it is also the necessary condition it is sufficient to show that the function $V^S(A_2)$ is quasi-concave for all feasible values of A_2 . The function $V^S(A_2)$ has nice analytical properties, and numerical computations show that it is indeed the case. This suggests that for $x \in [\bar{y}_N, \underline{y}_N]$ and for $x \geq \bar{y}_4$, the sufficient condition for burning money, Proposition 2, is also the necessary condition, implying that $\Omega^S = \Omega^{CT}$ for all $x \in [\bar{y}_N, \underline{y}_N]$. We do not pursue the aim to prove this claim here.

Proof of Proposition 3.

Let $\mathbf{A}(\Omega)$ have a separating interval $(B_i, A_{i+1}) \in \mathbf{A}$. If a type $t \in (B_i, A_{i+1})$ signals as if it were of type $y \in (B_i, A_{i+1})$, he gets utility $u^S(t, y) = W(t, y) - b^S(y)$. Since Ω is a WPBE, $u^S(t, y)$ attains its maximum at $y = t$. The first-order condition $0 = \partial u^S / \partial y$ at $y = t$ yields

$$b^S(t) = \int_{B_i}^t W_a(y, y) dy + b_i^S.$$

Hence, types $t \in (B_i, A_{i+1})$ get the following utility u^Ω in Ω :

$$u^\Omega = u^S(t, t) = W(t, t) - b^S(t) = W(t, t) - \int_{B_i}^t W_a(y, y) dy - b_i^S.$$

Let us construct a WPBE $\tilde{\Omega}$ which only differs from Ω by that types $t \in (A_{i+1} - \varepsilon, A_{i+1}) \subset (B_i, A_{i+1})$ pool, i.e., $\mathbf{A}(\tilde{\Omega}) = \mathbf{A}(\Omega) \cup \{(A_{i+1} - \varepsilon, A_{i+1})\}$, and type $t = A_{i+1} - \varepsilon$ burns the same amount as in Ω , i.e., $b^\Omega(A_{i+1} - \varepsilon) = b^{\tilde{\Omega}}(A_{i+1} - \varepsilon)$. Expected type and, hence, optimal action $a(\varepsilon)$ for $t \in (A_{i+1} - \varepsilon, A_{i+1})$ is:

$$a(\varepsilon) \equiv E(t | t \in (A_{i+1} - \varepsilon, A_{i+1})) = \int_{A_{i+1} - \varepsilon}^{A_{i+1}} t dF / \int_{A_{i+1} - \varepsilon}^{A_{i+1}} dF.$$

Types $t \in (A_{i+1} - \varepsilon, A_{i+1})$ get utilities $\tilde{u}^P(t, t) = W(t, a(\varepsilon)) - b^P(\varepsilon)$ where $b^P(\varepsilon)$ is the amount of money burned in interval $(A_{i+1} - \varepsilon, A_{i+1})$ in $\tilde{\Omega}$. By construction, it must be that $\tilde{u}^P(t, t) = u^\Omega$ at $t = A_{i+1} - \varepsilon$, which determines $b^P(\varepsilon)$:

$$b^P(\varepsilon) = W(A_{i+1} - \varepsilon, a(\varepsilon)) - W(A_{i+1} - \varepsilon, A_{i+1} - \varepsilon) + \int_{B_i}^{A_{i+1} - \varepsilon} W_a(y, y) dy + b_i^S.$$

Therefore, types $t \in (A_{i+1} - \varepsilon, A_{i+1})$ get utility $\tilde{u}^P(t, t)$.

$$\tilde{u}^P(t, t) = W(t, a(\varepsilon)) - W(A_{i+1} - \varepsilon, a(\varepsilon)) + W(A_{i+1} - \varepsilon, A_{i+1} - \varepsilon) - \int_{B_i}^{A_{i+1} - \varepsilon} W_a(y, y) dy - b_i^S.$$

Type $t = A_{i+1}$ gets utility $u^{\tilde{\Omega}}(A_{i+1}) = \tilde{u}^P(A_{i+1}, A_{i+1})$ in $\tilde{\Omega}$ whereas it gets $u^\Omega(A_{i+1}) = u^S(A_{i+1}, A_{i+1})$ in Ω . We consider the cases $u^{\tilde{\Omega}}(A_{i+1}) \geq u^\Omega(A_{i+1})$ and $u^{\tilde{\Omega}}(A_{i+1}) \leq u^\Omega(A_{i+1})$ separately.

a) Let $u^{\tilde{\Omega}}(A_{i+1}) \leq u^\Omega(A_{i+1})$. In this case, when types $t \in (A_{i+1} - \varepsilon, A_{i+1})$ pool, all types

$t \geq A_{i+1}$ have to burn extra amount of money $\varphi(\varepsilon) \equiv u^\Omega(A_{i+1}) - u^{\tilde{\Omega}}(A_{i+1}) \geq 0$ relative to

what they burn in Ω . (explanation needed) Using the properties of the function $a(\varepsilon)$,

i.e., $a(0) = A_{i+1}$, $a'(0) = -0.5$, and $a''(0) = f'(A_{i+1}) / (6f(A_{i+1}))$, it is a routine to show that

$\varphi(0) = \varphi'(0) = \varphi''(0) = 0$ and

$$\varphi'''(0) = \frac{W_{ta}(A_{i+1}, A_{i+1})}{2} \left(\Psi(A_{i+1}) - \frac{f'(A_{i+1})}{f(A_{i+1})} \right).$$

This implies that if $\Psi > f' / f$ at $t = A_{i+1}$, there is an $\bar{\varepsilon} > 0$ such that $\varphi(\varepsilon) > 0$ for all $\varepsilon \in (0, \bar{\varepsilon})$. The net *ex-ante* expected utility gain from pooling types $t \in (A_{i+1} - \varepsilon, A_{i+1})$ is:

$$\begin{aligned} g(\varepsilon) &\equiv E(u^{\tilde{\Omega}}(t)) - E(u^{\Omega}(t)) \\ &= \int_{A_{i+1}-\varepsilon}^{A_{i+1}} (u^P(t, t) - u^S(t, t)) dF + (u^P(A_{i+1}, A_{i+1}) - u^S(A_{i+1}, A_{i+1}))(1 - F(A_i)) \end{aligned}$$

It is a routine to show that $g(0) = g'(0) = g''(0) = 0$, and

$$g'''(0) = \frac{(1-F)}{2} W_{ia} \left(\frac{f}{(1-F)} + \frac{f'}{f} - \Psi \right).$$

at $t = A_{i+1}$. Thus, when

$$\frac{f'}{f} < \Psi < \frac{f}{(1-F)} + \frac{f'}{f},$$

the Sender strictly prefers WPBE $\tilde{\Omega}$ to WPBE Ω .

- b) Let $u^{\tilde{\Omega}}(A_{i+1}) \geq u^{\Omega}(A_{i+1})$. In this case, when types $t \in (A_{i+1} - \varepsilon, A_{i+1})$ pool all types $t \geq A_{i+1}$ must burn less money, which might not be feasible. However, we may require instead that all types $t \leq A_{i+1} - \varepsilon$ burn an extra amount of money $-\varphi(\varepsilon) \geq 0$ relative to what they burn in Ω . In a similar fashion, it can be shown that when

$$\frac{f'}{f} - \frac{f}{F} < \Psi < \frac{f'}{f},$$

the Sender strictly prefers WPBE $\tilde{\Omega}$ to WPBE Ω .

The only special case left is $\Psi(t) = f' / f$. When this happens, whether $u^{\tilde{\Omega}}(A_{i+1}) \geq u^{\Omega}(A_{i+1})$ or $u^{\tilde{\Omega}}(A_{i+1}) \leq u^{\Omega}(A_{i+1})$ is determined by higher-order derivatives. Nevertheless, one of these inequalities will hold and, therefore, conditions $-f / F < [\Psi - f' / f] < f / (1 - F)$ will hold due to the assumption $f > 0$. Q.E.D.

Proof of Proposition 4.

The proof is done by contradiction. Suppose that condition (9) fails for all $t \in (0,1)$. This would imply that $[\Psi - f' / f] > f / (1 - F) > 0$ for $t = 0$ and that $0 > -f / F > [\Psi - f' / f]$ for $t = 1$. Since $\Psi - f' / f$ is continuous, there exists a $\hat{t} \in (0,1)$ such that $\Psi - f' / f = 0$ at $t = \hat{t}$ and, therefore, condition (9) must hold at $t = \hat{t}$, the desired contradiction. *Q.E.D.*

Inequalities which proofs are omitted in the main text.

1. Proof of Lemma 3.

Show that $\hat{V}^S(A_k)$ is convex, where

$$\hat{V}^S(A_k) = -c - x^2 - x + \frac{1}{6} \left(\sum_{i=1}^{k-n-2} D_i^3 + D_{k-n-1}^3 + \sum_{j=1}^n D_{k-j}^3 + D_k^3 + \sum_{i=k+1}^N D_i^3 \right).$$

We compute the first and the second-order derivatives:

$$\begin{aligned} \frac{d\hat{V}^S}{dA_k} &= \frac{1}{6} \frac{d}{dA_k} \left(\underbrace{\sum_{i=1}^{k-n-2} D_i^3 + \sum_{i=k+1}^N D_i^3}_{=0} \right) + \frac{1}{6} \frac{d}{dA_k} \left(D_{k-n-1}^3 + \sum_{j=1}^n D_{k-j}^3 + D_k^3 \right) \\ &= \frac{1}{6} \frac{d}{dA_k} \left(D_{k-n-1}^3 + 4x(A_k^2 - A_{k-n}^2) + 4c(A_k - A_{k-n}) + D_k^3 \right) \\ &= \frac{1}{6} \left((3D_{k-n-1}^2 - 8xA_{k-n} - 4c) \frac{dA_{k-n}}{dA_k} + 8xA_k + 4c - 3D_k^2 \right), \end{aligned}$$

and

$$\begin{aligned} \frac{d^2\hat{V}^S}{dA_k^2} &= \left(D_{k-n-1} \left(\frac{dA_{k-n}}{dA_k} \right)^2 + \frac{1}{2} D_{k-n-1}^2 \frac{d^2A_{k-n}}{dA_k^2} + D_k \right) \\ &\quad + \frac{1}{6} \left(8x \left(1 - \left(\frac{dA_{k-n}}{dA_k} \right)^2 \right) - 4(2xA_{k-n} + c) \frac{d^2A_{k-n}}{dA_k^2} \right). \end{aligned} \tag{17}$$

Evaluating dA_{k-n}/dA_k and d^2A_{k-n}/dA_k^2 yields:

$$\frac{dA_{k-n}}{dA_k} = 1 - 2n \frac{ds(A_k, c)}{dA_k} = 1 - \frac{2xn}{s(A_k, c)}, \text{ and } \frac{d^2A_{k-n}}{dA_k^2} = \frac{2xn}{s^2(A_k, c)} \frac{ds(A_k, c)}{dA_k} = \frac{2x^2n}{s^3(A_k, c)} > 0.$$

Hence, the first bracket in (17) is strictly positive, so that:

$$\begin{aligned} \frac{d^2\hat{V}^S}{dA_k^2} &> \frac{1}{6} \left(8x \left(1 - \left(\frac{dA_{k-n}}{dA_k} \right)^2 \right) - 4(2xA_{k-n} + c) \frac{d^2A_{k-n}}{dA_k^2} \right) \\ &= \frac{4x^2n(4s(A_k, c)(s(A_k, c) - xn) - 2xA_{k-n} - c)}{3s^3(A_k, c)} = \frac{4x^2n(3s^2(A_k, c) - (4n^2 - 1)x^2)}{3s^3(A_k, c)}. \end{aligned}$$

Finally, $D_{k-n} > 0$ implies $s(A_k, c) > (2n - 1)x$, which yields:

$$\frac{d^2 \hat{V}^S}{dA_k^2} > \frac{4x^4 n}{3s^3(A_k, c)} (3(2n-1)^2 - (4n^2 - 1)) = \frac{16x^4 n(n-1)(2n-1)}{3s^3(A_k, c)} \geq 0.$$

Hence, $\hat{V}^S(A_k)$ is strictly convex.

2. Proof of Lemma 4, part 1.

Show that $dV^S / dA_N > 0$, where

$$\frac{dV^S}{dA_N}(A_N) = \frac{(N-1)}{6} \left(3A_2^2 - 8xA_2 - (1-A_N)^2 + 4(1+A_N)x - \frac{(2(1-A_2)-3)}{(N-1)}((1-A_N)+2x) \right).$$

Using (15), we express A_N in terms of A_2 :

$$A_N = \frac{A_2}{(N-1)} - 2(N-2)x + \frac{(N-2)}{(N-1)}.$$

Then, we express dV^S / dA_N in terms of A_2 :

$$\frac{dV^S}{dA_N} = \frac{(N-1)}{6} \left(\frac{3A_2(1+A_2(N-2)N)}{(N-1)^2} + 2x(3-2x(N-2)N) \right).$$

Next, we define \hat{A} by $c^P(\hat{A}, A_2, c) = 0$:

$$\hat{A} \equiv 1 - N((1-A_N) - 2(N-1)x) = 1 - N\left(\frac{(1-A_2)}{(N-1)} - 2x\right).$$

Since (4) implies $\partial c^P / \partial A = x + (B-A)/2 > 0$, and $c_1^P = c^P(0, A_2, c) > 0$ by our assumption, it

must be that $\hat{A} < 0$, which is equivalent to:

$$2Nx < (1-A_2)N/(N-1) - 1.$$

Hence,

$$3 - 2x(N-2)N > 3 - (N-2)\left(\frac{(1-A_2)N}{(N-1)} - 1\right) = \frac{(2N-1) + (N-2)NA_2}{(N-1)} > 0,$$

and, therefore, $dV^S / dA_N > 0$. This ends the proof of part 1.

3. Proof of Lemma 4, part 2.

Show that $\tilde{A}_{k+1} \in (A_k, A_{k+1})$ and $\tilde{c}_k^P = c^P(A_k, \tilde{A}_{k+1}, c) > 0$, where

$$\tilde{A}_{k+1} \equiv A_{k+2} - 2\left(\sqrt{x^2 + 2xA_{k+2} + c} - x\right).$$

We fix the value of A_{k+1} , define $v \equiv A_{k+2} - A_{k+1} > 0$, and formally consider A_{k+2} as a function of v , *i.e.*, $A_{k+2}(v) = A_{k+1} + v$. Similarly, we consider functions

$$\tilde{A}_{k+1}(v) = A_{k+2}(v) - 2\left(\sqrt{x^2 + 2xA_{k+2}(v) + c} - x\right) \text{ and}$$

$$\tilde{D}_{k+1}(v) \equiv A_{k+2}(v) - \tilde{A}_{k+1}(v) = 2\left(\sqrt{x^2 + 2xA_{k+2}(v) + c} - x\right).$$

First, $\tilde{A}_{k+1}(0) = A_k$ since $c^P(A_k, A_{k+1}, c) = 0$. Then, $\tilde{A}_{k+1}(\hat{A}_{k+2} - A_{k+1}) = A_{k+1}$ since $c^P(A_{k+1}, \hat{A}_{k+2}, c) = 0$. Finally, $\tilde{D}_{k+1}(0) = 4kx + A_2$. The derivatives are

$$\frac{d\tilde{A}_{k+1}}{dv} = 1 - \frac{2x}{\sqrt{x^2 + 2xA_{k+2} + c}} = \frac{\tilde{D}_{k+1} - 2x}{\tilde{D}_{k+1} + 2x}, \text{ and}$$

$$\frac{d\tilde{D}_{k+1}}{dv} = 1 - \frac{d\tilde{A}_{k+1}}{dv} = \frac{4x}{\tilde{D}_{k+1} + 2x} > 0.$$

For $k \geq 2$, this implies that $\tilde{D}_{k+1}(0) \geq 2x$, and consequently $\tilde{D}_{k+1} > 2x$, $d\tilde{A}_{k+1}/dv > 0$, and finally $\tilde{A}_{k+1} \in (A_k, A_{k+1})$. Hence, $\tilde{c}_k^P = c^P(A_k, \tilde{A}_{k+1}, c) > 0$ and $\tilde{\Omega}$ is indeed WPBE.

4. Proof of Proposition 2, definitions of \underline{A}_2 and \bar{A}_2 .

Show that $A_3(A_2) > A_2$ and $c_2^P = c^P(A_2, A_3, c) \geq 0$ together imply $\bar{A}_2 > A_2 \geq \underline{A}_2$.

We define \tilde{A}_2 by $c^P(\tilde{A}_2, A_3, c) = 0$. Since $\partial c^P(A, B, c)/\partial A > 0$ holds, $c^P(A_2, A_3, c) \geq 0$ requires $A_2 \geq \tilde{A}_2$. Hence, we have

$$A_3(A_2) > A_2 \geq \tilde{A}(A_2),$$

where $\tilde{A}(A_2) = 1 - 2(N-1)(s(A_2) - (N-1)x)$. Then, we proceed:

$$\begin{cases} 1 - 2(N-2)(s(A_2) - (N-2)x) > A_2 \\ A_2 \geq 1 - 2(N-1)(s(A_2) - (N-1)x) \end{cases},$$

$$\begin{cases} 1 - A_2 > 2(N-2)(s(A_2) - (N-2)x) \\ 2(N-1)(s(A_2) - (N-1)x) \geq 1 - A_2 \end{cases},$$

and, finally,

$$\frac{1-A_2}{2(N-2)} + (N-2)x > s(A_2) \geq \frac{1-A_2}{2(N-1)} + (N-1)x$$

Using

$$s(A_2) = \frac{1}{2} \sqrt{(2x - A_2)^2 + 8x},$$

we continue:

$$\begin{cases} \frac{(1-A_2)^2}{(N-2)^2} + 4x(1-A_2) + 4(N-2)^2 x^2 > (2x - A_2)^2 + 8x \\ (2x - A_2)^2 + 8x \geq \frac{(1-A_2)^2}{(N-1)^2} + 4x(1-A_2) + 4(N-1)^2 x^2 \end{cases}$$

It is routine to show that the above system can be written as follows:

$$\begin{cases} \left(A_2 + \frac{1-2(N-2)(N-3)x}{(N-3)} \right) \left(A_2 - \frac{1-2(N-1)(N-2)x}{(N-1)} \right) < 0 \\ \left(A_2 + \frac{1-2(N-1)(N-2)x}{(N-2)} \right) \left(A_2 - \frac{1-2N(N-1)x}{N} \right) \geq 0 \end{cases},$$

or

$$\begin{cases} (A_2 + 2(N-2)(x_{N-2} - x))(A_2 - 2(N-2)(x_{N-1} - x)) < 0 \\ (A_2 + 2(N-1)(x_{N-1} - x))(A_2 - 2(N-1)(x_N - x)) \geq 0 \end{cases}$$

Here we consider 4 cases by assigning different signs to all four multipliers.

Consider 4 cases:

Case 1.

$$\begin{cases} (A_2 + 2(N-2)(x_{N-2} - x)) > 0 \\ (A_2 - 2(N-2)(x_{N-1} - x)) < 0 \\ (A_2 + 2(N-1)(x_{N-1} - x)) \geq 0 \\ (A_2 - 2(N-1)(x_N - x)) \geq 0 \end{cases},$$

which reduces to

$$\begin{cases} A_2 < 2(N-2)(x_{N-1} - x) = \overline{A_2} \\ A_2 \geq 2(N-1)(x_N - x) = \underline{A_2} \end{cases},$$

as desired. Now we will show that the other three cases can never realize.

Case 2.

$$\begin{cases} (A_2 + 2(N-2)(x_{N-2} - x)) < 0 \\ (A_2 - 2(N-2)(x_{N-1} - x)) > 0 \\ (A_2 + 2(N-1)(x_{N-1} - x)) \geq 0 \\ (A_2 - 2(N-1)(x_N - x)) \geq 0 \end{cases},$$

$$\begin{cases} A_2 < 2(N-2)(x - x_{N-2}) \\ A_2 > 2(N-2)(x_{N-1} - x) \\ A_2 \geq 2(N-1)(x - x_{N-1}) \\ A_2 \geq 2(N-1)(x_N - x) \end{cases},$$

from which it follows that $2(N-2)(x - x_{N-2}) \geq 2(N-1)(x - x_{N-1})$, and, finally,

$$x \leq (N-1)x_{N-1} - (N-2)x_{N-2} = -\frac{1}{2(N-2)(N-3)} < 0.$$

This case can never happen.

Case 3.

$$\begin{cases} (A_2 + 2(N-2)(x_{N-2} - x)) > 0 \\ (A_2 - 2(N-2)(x_{N-1} - x)) < 0 \\ (A_2 + 2(N-1)(x_{N-1} - x)) \leq 0 \\ (A_2 - 2(N-1)(x_N - x)) \leq 0 \end{cases},$$

from which it follows that

$$\begin{cases} A_2 < 2(N-2)(x_{N-1} - x) \\ A_2 \leq 2(N-1)(x - x_{N-1}) \end{cases},$$

and, finally,

$$\begin{cases} x_{N-1} - x > 0 \\ x - x_{N-1} \geq 0 \end{cases}.$$

This case can never happen.

Case 4.

$$\begin{cases} (A_2 + 2(N-2)(x_{N-2} - x)) < 0 \\ (A_2 - 2(N-2)(x_{N-1} - x)) > 0 \\ (A_2 + 2(N-1)(x_{N-1} - x)) \leq 0 \\ (A_2 - 2(N-1)(x_N - x)) \leq 0 \end{cases},$$

$$\begin{cases} A_2 < 2(N-2)(x - x_{N-2}) \\ A_2 > 2(N-2)(x_{N-1} - x) \\ A_2 \leq 2(N-1)(x - x_{N-1}) \\ A_2 \leq 2(N-1)(x_N - x) \end{cases},$$

from which it follows that $2(N-2)(x_{N-1} - x) \leq 2(N-1)(x_N - x)$, and, finally,

$$x \leq (N-1)x_N - (N-2)x_{N-1} = -\frac{1}{2N(N-1)} < 0,$$

This case can never happen.

Thus, $A_3(A_2) > A_2$ and $c_2^P = c^P(A_2, A_3, c) \geq 0$ imply $\bar{A}_2 > A_2 \geq \underline{A}_2$.

5. Proof of Proposition 2, Part1, inequality $G(A_2) > 0$ (done numerically).

Show that $G(A_2) > 0$ for all $A_2 \in [0, \bar{A}_2)$, and $x \in [\hat{x}_{N-1}, x_{N-1})$, where

$$G(A_2) = \frac{1}{6} \left(A_3(4(A_2 + x) - 3A_3) - A_2 + 2x + (3(A_3 - A_2)^2 - A_2^2 - 4x(2A_3 - A_2)) \frac{dA_3}{dA_2}(A_2) \right)$$

Analytical pre-requisites:

$$\bar{A}_2 = 2(N-2)(x_{N-1} - x) = \frac{1}{(N-1)} - 2(N-2)x$$

$$2x - A_2 > 2x - 2(N-2)(x_{N-1} - x) = 2(N-1)(x - \hat{x}_{N-1}) \geq 0$$

$$s(A_2) = \frac{1}{2} \sqrt{(2x - A_2)^2 + 8x} > \frac{1}{2} \sqrt{(2x - \bar{A}_2)^2 + 8x} = (N-1)x + \frac{1}{2(N-1)}$$

$$A_3(A_2) = 1 - 2(N-2)(s(A_2) - (N-2)x) < \frac{1}{(N-1)} - 2(N-2)x$$

$$\frac{dA_3}{dA_2}(A_2) = \frac{(N-2)(2x - A_2)}{2s} < \frac{(N-1)(N-2)2x}{2(N-1)^2 x + 1} < (N-1)(N-2)x$$

6. Proof of Proposition 2, Part 1, inequality $G(\underline{A}_2) > 0 > G(\bar{A}_2)$.

Show that $G(\underline{A}_2) > 0 > G(\bar{A}_2)$ at $x = x_N$.

At $A_2 = \bar{A}_2$:

$$\begin{aligned}
x = x_N &= \frac{1}{2N(N-1)} \\
\bar{A}_2 &= 2(N-2)(x_{N-1} - x) = \frac{1}{(N-1)} - 2(N-2)x = \frac{2}{N(N-1)} \\
A_3(\bar{A}_2) &= \frac{1}{(N-1)} - 2(N-2)x = \bar{A}_2 = \frac{2}{N(N-1)} \\
s(\bar{A}_2) &= (N-1)x + \frac{1}{2(N-1)} = \frac{(2N-1)}{2N(N-1)} \\
A_3(\bar{A}_2) &= \frac{1}{(N-1)} - 2(N-2)x = \bar{A}_2 = \frac{2}{N(N-1)} \\
\frac{dA_3}{dA_2}(\bar{A}_2) &= \frac{(N-2)(2x - \bar{A}_2)}{2s} = \frac{(N-2)\left(\frac{1}{(N-1)} - 2(N-1)x\right)}{2s} = \frac{(N-2)}{(2N-1)}, \\
G(\bar{A}_2) &= \frac{1}{6} \left(A_3(4(A_2 + x) - 3A_3) - A_2 + 2x + (3(A_3 - A_2)^2 - A_2^2 - 4x(2A_3 - A_2)) \frac{dA_3}{dA_2} \right) \\
&= -\frac{1}{6} \left(\bar{A}_2 \left(1 - 4x + (\bar{A}_2 + 4x) \frac{dA_3}{dA_2} - \bar{A}_2 \right) - 2x \right) \\
&= -\frac{1}{6} \left(\bar{A}_2 \left(1 + 2(N-4)x + \left(\frac{1}{(N-1)} - 2(N-4)x \right) \frac{dA_3}{dA_2} - \frac{1}{(N-1)} \right) - 2x \right) \\
&= -\frac{1}{6N(N-1)(N-1)(2N-1)} \left((2N+1)(N-3) + 2 + \frac{2(N-4)(N+1)}{N} \right) < 0.
\end{aligned}$$

At $A_2 = \underline{A}_2$:

$$\begin{aligned}
x = x_N &= \frac{1}{2N(N-1)} \\
\underline{A}_2 &= 2(N-1)(x_N - x) = 0 \\
s(0) &= \sqrt{x^2 + 2x} = x \sqrt{1 + \frac{2}{x}} = \frac{(2N-1)}{2N(N-1)} \\
A_3(A_2) &= \frac{2}{N(N-1)} \\
\frac{dA_3}{dA_2} &= \frac{(N-2)(2x - A_2)}{2s} = \frac{(N-2)}{(2N-1)}
\end{aligned}$$

$$\begin{aligned}
G(0) &= \frac{1}{6} \left(A_3(4x - 3A_3) + 2x + A_3(3A_3 - 8x) \frac{dA_3}{dA_2} \right) \\
&= \frac{N(N-1)(2N-1) + (4N-8) - 8(2N-1)}{6N^2(N-1)^2(2N-1)} = \frac{((N-4)(2N+5)+9)N}{6N^2(N-1)^2(2N-1)} > 0.
\end{aligned}$$

Thus, $G(\underline{A}_2) > 0 > G(\bar{A}_2)$ at $x = x_N$.

7. Proof of Proposition 2, Part 1, inequality $G(\underline{A}_2) < 0$.

Show that $G(\underline{A}_2) < 0$ at $x = \hat{x}_N$.

At $A_2 = \underline{A}_2$:

$$\begin{aligned}
x &= \hat{x}_N = \frac{1}{2N^2} \\
x_N &= \frac{1}{2N(N-1)} \\
\underline{A}_2 &= 2(N-1)(x_N - x) = \frac{1}{N^2} \\
s(\underline{A}_2) &= \frac{1}{N} \\
A_3(\underline{A}_2) &= 1 - 2(N-2)(s - (N-2)x) = \frac{4}{N^2} \\
\frac{dA_3}{dA_2} &= \frac{(N-2)(2x - A_2)}{2s} = 0 \\
G(\underline{A}_2) &= \frac{1}{6} (A_3(4(A_2 + x) - 3A_3) - A_2 + 2x) = -\frac{4}{N^4} < 0
\end{aligned}$$

Thus, $G(\underline{A}_2) < 0$ at $x = \hat{x}_N$.

8. Proof of Proposition 2, Part 1, inequality $G(A_2) < 0$ (done numerically).

Show that $G(A_2) < 0$ when $c(A_2) > 0$, $x \in [\hat{x}_N, x_N)$, and $A_2 \in [\underline{A}_2, \bar{A}_2)$.

9. Proof of Proposition 2, Part 1, inequality $G(A_2) < 0$ (done numerically).

Show that $G(A_2) < 0$ for all $A_2 \in [\underline{A}_2, \bar{A}_2)$ and $x < x_{N+1}$.

10. Proof of Proposition 2, Part 2, inequality $G(0) > 0 > G(\bar{A}_2)$.

Show that $G(0) > 0 > G(\bar{A}_2)$ at $x = x_N$. The proof is exactly the same as the proof of $G(\underline{A}_2) > 0 > G(\bar{A}_2)$ at $x = x_N$ above.

11. Proof of Proposition 2, Part 2, inequality $G(\bar{A}_2) > 0$.

Show that $G(\bar{A}_2) > 0$ at $x = \hat{x}_{N-1}$.

$$\begin{aligned}
x &= \hat{x}_{N-1} = \frac{1}{2(N-1)^2} \\
\bar{A}_2 &= 2(N-2)(x_{N-1} - x) = \frac{1}{(N-1)^2} \\
s(\bar{A}_2) &= \frac{1}{2} \sqrt{(2x - A_2)^2 + 8x} = \frac{1}{(N-1)} \\
A_3(A_2) &= 1 - 2(N-2)(s - (N-2)x) = \frac{1}{(N-1)^2} \\
\frac{dA_3}{dA_2}(\bar{A}_2) &= \frac{(N-2)(2x - \bar{A}_2)}{2s} = 0 \\
G(\bar{A}_2) &= \frac{1}{6} \left(A_3(4(A_2 + x) - 3A_3) - A_2 + 2x + (3(A_3 - A_2)^2 - A_2^2 - 4x(2A_3 - A_2)) \frac{dA_3}{dA_2} \right) \\
&= \frac{1}{6} (A_3(4(A_2 + x) - 3A_3) - A_2 + 2x) = \frac{1}{2(N-1)^4} > 0
\end{aligned}$$

12. Proof of Proposition 2, Part 2, inequality $G(A_2) < 0$ (done numerically).

Show that $G(A_2) < 0$ for $A_2 \in [0, \bar{A}_2)$ and $x < \hat{x}_N$.